JOURNAL OF MEMORY AND LANGUAGE **37,** 382–410 (1997)
ARTICLE NO. ML972516

Orthographic and Phonological Neighborhoods in Naming: Not All Neighbors Are Equally Influential in Orthographic Space

Ronald Peereman

Laboratoire d'Etude des Apprentissages et du Developpement, CNRS, ESA 5022, Université de Bourgogne, Dijon, France

and

Alain Content

Laboratoire de Psychologie Expérimentale, Université Libre de Bruxelles, Brussels, Belgium

The neighborhood size effect refers to the finding that single word naming is faster for stimuli that are orthographically similar to numerous lexical entries. We explored the nature of this phenomenon in five experiments with French pseudowords and words, and we examined the orthographic and the phonological characteristics of neighbors through quantitative analyses of a word corpus. Experiments 1 and 2 showed that the facilitatory effect of neighborhood size was determined by a subset of neighbors, called *phonographic* neighbors, which are also phonologically similar to the target letter string. Experiments 3 to 5 aimed at assessing the influence of phonographic neighbors as a function of the constituents shared with the target. The results suggested that the number of neighbors sharing the target rime determines the facilitation effect. The findings are discussed in relation to the structure of the French orthography and its characteristics in comparison with English. We conclude that the joined orthographic and phonological similarity between lexical neighbors and the target letter string determines the facilitation effect observed in naming.

During the past few years, much evidence has shown that reading performance is influenced by the orthographic similarity between the letter string being processed and other words. This evidence caused changes in serial models of lexical access (Forster, 1987) and models of print-to-sound conversion (Coltheart, Curtis, Atkins, & Haller, 1993). The pool of words orthographically related to the target letter string is usually called the neighborhood of the word.

In the naming task, performance is improved when words or pseudowords have many orthographic neighbors (Andrews, 1989, 1992; Laxon, Masterson, & Moran, 1994; Laxon, Masterson, Pool, & Keating, 1992; McCann & Besner, 1987; Peereman & Content, 1995), the neighborhood size (N size) effect for words being essentially confined to low-frequency items (Andrews, 1989; Peereman & Content, 1995; Sears, Hino, & Lupker, 1995). However, there have been few attempts to investigate the functional locus of the effect.

The approach adopted here is to examine

This work was supported by the French CNRS, by a grant from the Direction générale de la Recherche scientifique—Communauté française de Belgique (ARC Grants 91/96-148 and 96/01-203), and by a joint French–Belgian grant for scientific exchanges (Programme Tournesol). We thank Marielle Lange and Josiane Lechat for their help in conducting Experiment 3b. We are grateful to two anonymous reviewers for their useful comments and constructive criticisms on previous versions.

Address correspondence and reprint requests to Ronald Peereman, Laboratoire d'Etude des Apprentissages et du Développement, CNRS, ESA 5022, Université de Bourgogne, Faculté des Sciences, 6 boulevard Gabriel, F-21000 Dijon, France. E-mail: peereman@satie.u-bourgogne.fr.

0749-596X/97 $25.00

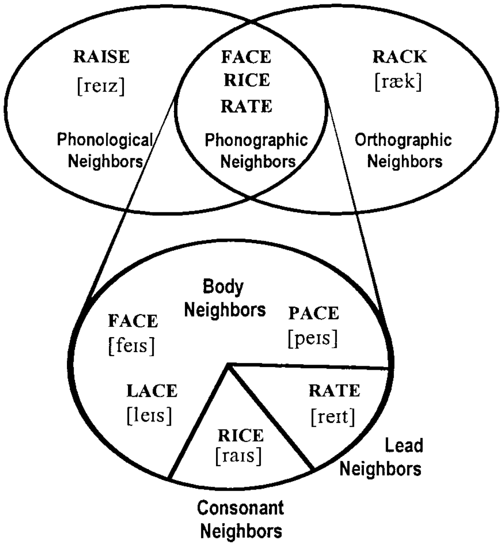

FIG. 1. An example illustrating the different types of lexical neighbors considered in the present study, for the English word RACE [**reIS].**

the contribution of different types of neighbors defined along both the orthographic and the phonological dimensions. There are two reasons to consider the phonological characteristics of the neighbors along with their orthographic properties. First, the naming task requires phonological encoding of the letter string. Second, because of the statistical regularity of the mapping between orthographic and phonological units in alphabetic writing systems, variations in orthographic neighborhood size entail correlated variations both in the frequency of the phonological units and in the frequency of the correspondences between orthographic and phonological patterns.

To help describe the analysis of neighborhood on which this study is based, let us consider the case of the word RACE depicted in the Venn diagrams of Fig. 1. The two ensembles at the top of Fig. 1 correspond to the set of orthographic neighbors and the set of phonological neighbors. Their intersection represents the pool of words that are both orthographically and phonologically similar to the target RACE. We refer to this subset as the *phonographic neighborhood.*

Following Coltheart, Davelaar, Jonasson, and Besner (1977), orthographic neighbors were operationally defined as the words of identical length (in number of letters) which can be generated by a single letter substitution. This definition was transposed to analyze phonological neighborhood, a phonological neighbor being any word of identical length (in phonemes) generated by a single phoneme substitution. Accordingly, the word RACK is an orthographic neighbor of RACE, but its phonological form differs by more than one phoneme from RACE. Thus, RACK is an orthographic neighbor but not a phonological neighbor. The opposite is true for the word RAISE. Its spelling differs by more than one letter from the word RACE, but its phonological form diverges by only one phoneme. Hence, the word RAISE is a phonological neighbor but not an orthographic neighbor of the word RACE. Finally, the words FACE, RICE, and RATE are phonographic neighbors.

The depiction of the phonographic neighborhood has been enlarged at the bottom of Fig. 1 to display a finer description distinguishing between three kinds of phonographic neighbors. These can diverge from the base word by the initial consonant, by the vowel, or by the final consonant. Thus, the target word and its phonographic neighbors can share the consonantal skeleton as in RICE, the vowel and the final consonant as in FACE or LACE, or the initial consonant and the vowel, as in RATE. We refer to these three types of phonographic neighbors as *consonant neighbors, body neighbors,* and *lead neighbors,* respectively.

Variations in the number of orthographic neighbors must be associated to variations in the number of phonographic neighbors. Also, an increase in the number of orthographic neighbors should be accompanied by an increase in the number of phonological neighbors. This relation is illustrated in the scattergram of Fig. 2. It plots the number of phonological neighbors as a function of the number of orthographic neighbors, for the

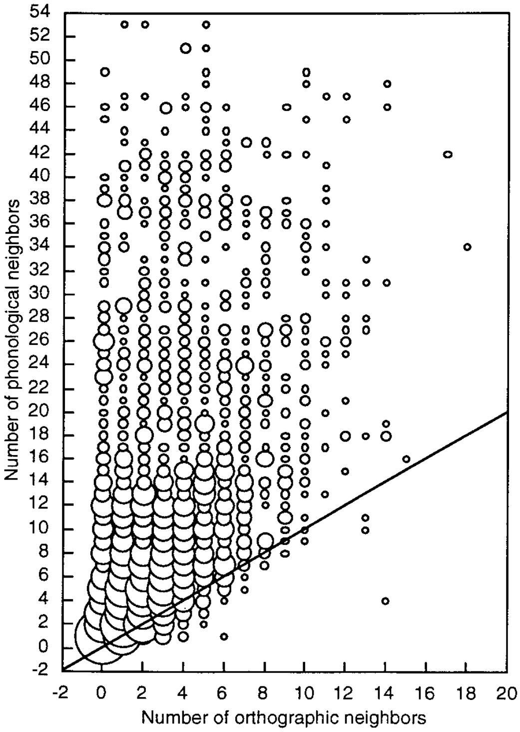

FIG. 2. Relation between orthographic neighborhood size and phonological neighborhood size. The area of each circle is proportional to the number of words represented.

2004 French monosyllabic four- to six-letter words appearing in the computerized lexical database Brulex (Content, Mousty, & Radeau, 1990). As expected, the number of phonological neighbors increases with orthographic neighborhood size. Hence, in view of the natural correlation between orthographic, phonological and phonographic neighborhood density, a functional interpretation of the performance variations associated with lexical neighborhoods requires disentangling the respective contributions of the different neighborhood categories. We shall examine each of the three possibilities in the context of the naming task.

A purely *orthographic* explanation of the neighborhood size effect holds that the number of orthographic neighbors is the chief neighborhood characteristic determining reading performance. Such a hypothesis was put forward by Andrews (1989, 1992) within the framework of the interactive activation model developed by McClelland and Rumelhart (1981). In this model, activation of the word nodes reverberates at the letter layer through feedback connections, therefore boosting the activation of the letter nodes. This in turn facilitates the orthographic encoding of the letter string and consequently accelerates lexical access. Because the amount of feedback activation is related to the number of word units activated, words that have many neighbors should be orthographically encoded faster than words with few neighbors.

A second proposal is that the benefit is caused by the *phonographic* neighbors, which facilitate phonological computation. This hypothesis received indirect support from studies of grapho-phonological consistency. Although those studies are generally based on a different conception of neighborhood, they indicate that naming is influenced by the ratio of the number of words in which the shared orthographic units are pronounced differently to the number of words in which they are pronounced similarly (Glushko, 1979; Jared, McRae, & Seidenberg, 1990; Laxon, Masterson, & Coltheart, 1991; Peereman, 1995). The account of the effect differs according to the theoretical framework considered. One class of models assumes that the analytical knowledge and the lexical knowledge activated by the letter string are combined in building a phonological code (e.g., Brown, 1987; Coltheart *et al.*, 1993; Shallice & McCarthy, 1985). So, lexical activation of convergent phonological information would facilitate phonological encoding. Another proposal belonging both to independent dual route models (Paap, McDonald, Schvaneveldt, & Noel, 1987; Patterson & Morton, 1985) and to the single process approach proposed by Seidenberg and McClelland (1989) is that the effect is an indirect consequence of the fact that letter strings with large orthographic neighborhoods generally include frequent grapho-phonological correspondences.

Finally, a third possibility is that the neighborhood size (N size) effect is due to the density of the *phonological* neighborhood. As shown in Fig. 2, there is a strong correlation between the size of the orthographic and the phonological neighborhoods. Phonological neighborhood size also correlates with phoneme frequency. Thus, as words from densely populated neighborhoods contain more frequent phonemes and phoneme groups, faster phonological and articulatory encoding might be expected.

The aim of the present study was to examine these different accounts of the N size effect in reading aloud. In Experiment 1, we assessed the phonological hypothesis, while in Experiment 2 we examined the influence of phonographic neighborhood. Quantitative analyses were performed to estimate the distributions of consonant, body, and lead neighbors as a function of orthographic neighborhood size. Given the emphasis on the body constituent in the English research, we also examined whether the structure of French orthographic and phonological word forms similarly favors the notion of an onset/rime decomposition. Finally, further experiments examined neighborhood effects as a function of the size of the different subsets of the phonographic neighborhood (body, lead, consonant skeleton).

Because it appeared difficult to match words on different characteristics such as word and bigram frequency, while at the same time manipulating separately orthographic and phonological neighborhoods, Experiments 1 to 4 used pseudowords. The main conclusions from the pseudoword studies were finally extended to words in Experiment 5. Matching the letter strings on the initial phoneme across pseudoword categories was not possible, and we therefore included a delayed naming task in each experiment to ensure that effects observed in the immediate naming task did not result from differences in articulation ease or in voice key sensitivity to the initial sounds of the stimuli.

EXPERIMENT 1: PHONOLOGICAL NEIGHBORHOOD

Experiment 1 examined whether the size of the phonological neighborhood influences pseudoword naming. The most obvious design would have been a factorial manipulation of the number of phonological and the number of orthographic neighbors. Unfortunately, as can be seen on Fig. 2, there are virtually no monosyllabic French words having fewer phonological neighbors than orthographic neighbors. Thus, only three categories of pseudowords were used. The first category included pseudowords with numerous orthographic neighbors and numerous phonological neighbors (ON+PN+). For example, the pseudoword VOULE /**vul**/ has 10 orthographic neighbors (e.g., BOULE, COULE, FOULE, VOILE, VEULE) and 18 phonological neighbors (e.g., BOULE /**bul**/, VOL /**vɔl**/, SAOULE /**sul**/, JOULE /**ʒul**/). The second category was made of pseudowords with few orthographic neighbors but numerous phonological neighbors (ON−PN+). For example, the pseudoword BAIME (**bɛm**/) has 17 phonological neighbors (e.g., BEC /**bɛk**/, BEIGE /**bɛʒ**/, THEME (**tɛm**/), but only one orthographic neighbor (BAUME). The pseudowords of the third category served as controls and had few orthographic and few phonological neighbors (ON−PN−). For example, the pseudoword FLIDE /**flid**/ has only one phonological neighbor (FLIC /**flik**/) and no orthographic neighbor.

If the size of the phonological neighborhood is the underlying factor controlling the N size effect, then pseudowords having many phonological neighbors should be pronounced faster than controls irrespective of the number of orthographic neighbors. In contrast, if the effect reflects the size of the orthographic neighborhood, faster naming should be expected for pseudowords having numerous orthographic neighbors than for pseudowords with few orthographic neighbors, irrespective of the size of the phonological neighborhood.

Method

Participants. The participants were 20 students at the University of Bourgogne. All were native speakers of French and received course credit for their participation.

TABLE 1

CHARACTERISTICS OF THE PSEUDOWORDS USED IN EXPERIMENT 1 (MEAN VALUES)

	ON+PN+		ON−PN+	
Characteristics	Tests	Controls	Tests	Controls
No. of neighbors/ON	10.0	0.8	0.8	0.2
Frequency of neighbors/ON[a]	1327.7	7.6	8.0	0.0
No. of neighbors/PN	22.9	1.3	21.0	2.1
Frequency of neighbors/PN[a]	2029.3	8.9	2460.1	21.7
No. of letters	4.8	4.8	4.8	4.8
Log bigram frequency[b]	3.2	3.1	3.0	2.9

Note. ON, orthographic neighborhood; PN, phonological neighborhood.
[a] Mean summed word frequency (per million) from Imbs (1971).
[b] Bigram frequencies were taken from Content and Radeau (1988).

Stimuli. The stimuli consisted of 80 monosyllabic pseudowords of four or five letters. Half had many phonological neighbors. Among these, 20 pseudowords had numerous orthographic neighbors (ON+PN+), and 20 had few orthographic neighbors (ON−PN+). The other half of the total set included pseudowords with few phonological neighbors and few orthographic neighbors (ON−PN−). Orthographic and phonological neighborhoods were computed on the French lexical database developed by Content et al. (1990). Descriptive statistics about the stimulus sets are shown in Table 1. The stimuli appear in Appendix A.

Procedure. The 80 experimental stimuli were divided in two blocks of identical length and the order of presentation of the two blocks was counterbalanced across subjects. The experimental session was preceded by 18 practice trials, and each block began with two warm-up trials. The stimuli were presented in lowercase on a computer screen. Presentation and timing were controlled by a PC286 connected with a voice key. An asterisk was presented for 200 ms at the center of the screen, followed by a 200-ms blank interval. Then the target letter string was presented in the center of the screen until the subject's response or 2 s elapsed. The intertrial interval was 2 s. The experimenter noted on separate protocols whether the pronunciation was incorrect and whether the recorded latency was invalid, in case of triggering of the voice key by extraneous sounds. Responses were considered as errors when they differed from the pronunciation based on grapheme–phoneme and body–rime correspondences. With only a few exceptions, there was little variability in the pronunciation of the pseudowords, reflecting the high consistency of French orthography.

In the immediate naming task, the subjects were told to read the pseudowords aloud as quickly and accurately as possible. After having completed the immediate naming task, the subjects performed a delayed naming task with the same lists of stimuli. In the delayed naming situation, the subjects were instructed to wait until the response cue appeared (a ''???'' sign) before pronouncing the letter string. The pseudoword was presented for 1500 ms and was followed by a blank screen. After a random delay interval of either 1300, 1400, or 1500 ms, the response cue was displayed and the time measured until the onset of the subject's response. To increase attention to the response cue, an auditory warning signal was presented 1 s after the removal of the letter string.

Results

For each experiment, separate analyses of variance of the immediate and delayed naming

TABLE 2

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF PSEUDOWORD CATEGORY IN EXPERIMENT 1

Pseudoword category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
ON+ PN+	528	3.3	334	0.5
ON− PN+	568	5.3	333	1.8
ON− PN−	565	7.8	338	2.6

Note. ON, orthographic neighborhood; PN, phonological neighborhood.

performances were carried out on Subjects and Items means (latencies and errors). Naming latencies smaller than 200 ms or longer than 1000 ms in the immediate naming task and smaller than 150 ms or longer than 900 ms in the delayed naming task were discarded, as were invalid response times and incorrect response times. Further analyses were systematically conducted on the differences between the immediate and delayed naming latencies. As they generally lead to the same conclusions as analyses on immediate latencies, they will be reported only when the outcomes are different.

In the present experiment, cutoffs led to the rejection of 0.3% of observations for immediate naming and 3.9% for the delayed naming task. Invalid latencies amounted to 3.2 and 3.1% in immediate and delayed naming, respectively.

A preliminary comparison indicated that there was no difference between the two sets of control pseudowords, and they were thus pooled together in the following analyses. The null difference between the two control sets is in line with previous reports that bigram frequency does not affect naming performance (Andrews, 1992; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). The mean naming latencies and error rates in immediate and delayed naming are presented in Table 2.

Note that two control pseudowords (TUNDE and JUNDE) were responsible for 31 and 52% of the errors in immediate and delayed naming, respectively. These two items were often pronounced /**tynd**/ and /ʒ**ynd**/ instead of /**tœ̃**d/ or /**ʒœ̃d**/ (as in the French pronunciation of the word JUNGLE /**ʒœ̃gl**/).

Immediate naming. A reliable effect of Pseudoword Category was obtained ($F_1(2, 38) = 41.98$, $p < .001$; $F_2(2,77) = 3.69$, $p < .05$). As can be seen in Table 2, ON+PN+ pseudowords gave rise to shorter naming latencies (40 ms) than ON−PN+ pseudowords ($F_1(1,38) = 67.68$, $p < .001$; $F_2(1,77) = 5.43$, $p < .025$). They were also pronounced 37 ms faster than ON−PN− pseudowords. This difference was reliable by Subjects and by Items ($F_1(1,38) = 57.87$, $p < .001$; $F_2(1,77) = 6.18$, $p < .025$). No difference was observed between the two sets of pseudowords that had few orthographic neighbors.

In the error rates analyses, the effect of Pseudoword Category was reliable by subjects, but far from significance by items ($F_1(1,38) = 18.35$, $p < .001$; $p = .26$ by items).

Delayed naming. No significant effect was observed on latencies. On the error rates, the effect of Pseudoword Category was significant in the analysis on subjects means only ($F_1(1,38) = 9.38$, $p < .025$; $p = .38$ by items).

Discussion

The main conclusion from Experiment 1 is that the purely phonological interpretation of neighborhood size effect is not supported. A large phonological neighborhood is not sufficient to produce any advantage relative to low neighborhood control pseudowords. A facilitation effect occurs only when the letter strings also have many orthographic neighbors. However, it remains unclear whether the orthographic neighborhood size constitutes the critical factor. The high orthographic neighborhood stimuli also had many phonological neighbors, and letter strings with large orthographic and phonological neighborhoods are likely to also have a large phonographic neighborhood. Indeed, the ON+PN+ pseudowords had, on average, 8.0 phonographic neighbors. Hence, the facilitation effect observed could be caused by the phonographic neighborhood. A previous study (Peereman & Content, 1995) provided preliminary evidence supporting this hypothesis. In a post hoc analysis, the orthographic neighbors of the target words were partitioned according to their phonological similarity with the stimuli. A multiple regression analysis showed that the number of phonologically close neighbors was a better predictor of naming latencies than the total number of neighbors. The role of the phonographic neighborhood was addressed in Experiment 2.

EXPERIMENT 2: PHONOGRAPHIC NEIGHBORHOOD

In Experiment 2, we examined whether the neighborhood size effect depends on the phonological properties of the orthographic neighbors. We contrasted two sets of pseudowords which had approximately the same number of orthographic neighbors in total, but differed by the proportion of phonographic neighbors: OPN+ pseudowords had many phonographic neighbors, whereas OPN− pseudowords had few phonographic neighbors. For example, the pseudoword VORTE /**vɔrt**ə/ has seven orthographic neighbors (VERTE /**vɛrt**ə/, FORTE /**fɔrt**ə/, VOLTE /**vɔlt**ə/, MORTE /**mɔrt**ə/, PORTE /**pɔrt**ə/, SORTE /**sɔrt**ə/, VOUTE /**vut**/), among which six are also phonological neighbors. Conversely, the pseudoword OURE /**ur**/ has 10 orthographic neighbors (BURE /**byr**/, CURE /**kyr**/, OCRE /ɔ**kr**ə/, DURE /**dyr**/, OGRE /ɔ**gr**ə/, PURE /**pyr**/, SURE /**syr**/, OURS /**urs**/, OUIE /**wi**/, HURE /**yr**/), but only the last one is also a phonological neighbor. To assess the presence of a facilitation effect, a third set of pseudowords having few orthographic and few phonological neighbors was used.

We should note that the classification of orthographic neighbors as phonological or nonphonological is not equivalent to the usual notion of consistency. First, phonological neighbors may or may not share the body/rime with the target (compare BORNE, CORNE, MORNE, and FORCE, FORGE, FORME, FORTE as neighbors of the pseudoword FORNE). Second, nonphonological neighbors are rarely exception words, given the scarcity of grapheme–phoneme irregularity in French. Most often, nonphonological neighbors are words for which a single letter substitution causes more than a single phoneme substitution in the phonemic representation. This happens essentially in the case of vowels represented by digraphs or in the case of consonants whose pronunciation is contextually determined by the following letter (as for the pronunciation of the letters G or C). For example, the words VIGIE (/**viʒi**/) and VIGNE (/**viɲ**/ are orthographic neighbors of the target pseudoword VIGLE (/**vigl**/), but none of them is a phonological neighbor. Similarly, the words FONTE (/**fɔ̃t**/) and FUITE (/**fɥit**/) are orthographic neighbors but not phonological neighbors of the pseudoword FOITE (/**fɥat**/).

Method

Participants. Twenty-nine students at the University of Bourgogne served as subjects for course credit.

Stimuli and Procedure. Three categories of French monosyllabic pseudowords three to five letters long were used (Appendix A). The first category (OPN+) included letter strings having a high proportion of phonographic neighbors. In the second category (OPN−), the orthographic neighbors of the pseudowords were mostly nonphonological neighbors. Finally, the pseudowords of the third category had few orthographic and few phonological neighbors. There were 18 items in each category. Descriptive statistics about the three stimulus sets appear in Table 3.

The pseudoword list started with two warm-up trials and was preceded by 18 practice trials. All subjects performed both the immediate and the delayed naming tasks—in fixed order—on the same stimuli. All other aspects of the procedure were identical to those in Experiment 1.

TABLE 3

CHARACTERISTICS OF THE PSEUDOWORDS USED IN EXPERIMENT 2 (MEAN VALUES)

Characteristics	OPN+	OPN−	Controls
No. of neighbors/ON	8.2	7.2	0.7
Frequency of neighbors/ON[a]	2046.0	1912.5	29.0
No. of neighbors/PN	22.3	12.8	4.0
Frequency of neighbors/PN[a]	5437.8	3135.7	77.8
No. of neighbors/OPN	7.2	2.1	0.2
Frequency of neighbors/OPN[a]	1986.7	42.7	7.8
No. of letters	4.1	4.1	4.1
Log bigram frequency	3.1	3.0	3.0

Note. ON, orthographic neighborhood; PN, phonological neighborhood; OPN, phonographic neighborhood.
[a] Mean summed word frequencies (per million) from Imbs (1971).

Results

Cutoffs led to the rejection of 0.1% of observations for immediate naming and 4.1% for delayed naming. Invalid latencies amounted to 1.7 and 3.4% in immediate and delayed naming, respectively. The mean naming latencies and the percentages of errors appear in Table 4.

Immediate naming. An analysis of variance comparing immediate naming latencies for the three pseudoword categories indicated a reliable effect ($F_1(2,56) = 73.07$, $MSE = 15{,}930$, $p < .001$; $F_2(2,51) = 5.68$, $MSE = 10{,}526$, $p < .01$). Latencies were 40 ms shorter for the OPN+ pseudowords than for the OPN-pseudowords ($F_1(1,56) = 104.92$, $MSE = 22{,}881$, $p < .001$; $F_2(1,51) = 7.53$, $MSE = 13{,}963$, $p < .01$). The 39-ms advantage of the OPN+ pseudowords relative to the Control pseudowords was also significant ($F_1(1,56) = 114.03$, $MSE = 24{,}868$, $p < .001$; $F_2(1,51) = 9.40$, $MSE = 17{,}424$, $p < .01$). The difference between OPN− pseudowords and control pseudowords was not significant.

The analyses of errors revealed a marginally significant effect of Pseudoword Category across subjects ($F_1(2,56) = 3.12$, $MSE = 2.172$, $p = .052$), but not across items ($p = .29$).

Delayed naming. There were no significant differences, either in the analysis on latencies ($F_1(2,56) = 1.71$, $MSE = 646$, $p = .19$; $p = .43$ by items), or in the analysis on errors ($p > .20$ in both analyses).

TABLE 4

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF PSEUDOWORD CATEGORY IN EXPERIMENT 2

Pseudoword category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
OPN+	513	2.7	323	0.4
OPN−	553	5.0	332	0.6
Controls	554	5.6	328	0.6

Note. OPN, phonographic neighborhood.

Discussion

Experiment 2 shows that pronunciation is facilitated only when the orthographic neighborhood includes many phonographic neighbors. Pseudowords with many orthographic neighbors but few phonographic neighbors were not pronounced faster than controls. Although Table 3 indicates that the two categories of pseudowords with many neighbors also differed on the number of phonological neighbors, the results of Experiment 1 lead us to reject an explanation based on the size of phonological neighborhood per se. Thus, taken together, the two experiments demonstrate that the effect is controlled by the size of the phonographic neighborhood.

Another interpretation would attribute lack of difference between the OPN− pseudowords and the controls to the cancellation of an orthographic facilitation effect by the interference of inconsistent neighbors. We shall return to this issue later. Regarding the present experiment, examination of the lists of neighbor words revealed that one pseudoword included an inconsistent body (TOS, where −OS can be pronounced either /ɔ/ as in most of the monosyllabic words—e.g., DOS, VOS, NOS—or /ɔ**s**/ as in the word OS and in most multisyllabic words—e.g., COSMOS, TETANOS). However, post hoc analyses indicated that the removal of this item did not alter the data pattern.

The data of Experiment 2 contradict the assumption that the *N* size effects in naming result from faster orthographic encoding. Hence, any account of the findings requires the incorporation of the phonological component specific to naming. In the next section, we explore the characteristics of the phonographic neighborhood and we examine the relation between neighborhood and consistency through lexical statistics.

ARE BODIES BUDDIES IN FRENCH? A DESCRIPTIVE ANALYSIS OF A LEXICAL CORPUS

Various findings in the English literature indicate that the final VC letter group of monosyllabic words, the *Body,* may have a special status in visual word recognition. Several authors have insisted on the early sensitivity to rhyme in prereading children (Bryant, MacLean, Bradley, & Crossland, 1990; Kirtley, Bryant, MacLean, & Bradley, 1989; Treiman, 1985) and the potential relevance of body–rime correspondences to the acquisition of phonological transcoding mechanisms (see Goswami & Bryant, 1990, for review). It has also been shown that body–rime consistency affects reading performance in children (Backman, Bruck, Hebert, & Seidenberg, 1984; Coltheart & Leahy, 1992; Laxon et al., 1991, 1994; Treiman et al., 1995) and adults (Andrews, 1982; Glushko, 1979, 1981; Jared et al., 1990; Kay & Bishop, 1987; Kay & Marcel, 1981; Laxon et al., 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1990; Treiman et al., 1995). Letter strings are named more slowly and less accurately when alternative phonological codes can be assigned to the body (e.g., -AVE).

The large variability of pronunciations of printed English vowels may render the print-to-sound conversion process sensitive to covariations between orthographic and phonological codes at the level of body–rime units. In a recent study, Treiman et al. (1995) showed that the reliability of vowel pronunciation in English CVC words greatly increased when the final consonants were also taken into account. No such improvement occurred when the consistency of onset + vowel pronunciation was estimated. In addition, the ease of phonological encoding may also be a function of the frequency of analytical print-to-sound associations. For instance, several studies have reported faster phonological conversion for frequent correspondences (e.g., Bowey & Hansen, 1994; Brown, 1987; Brown & Watson, 1994; Rosson, 1985; Treiman, Goswami, & Bruck, 1990; Treiman et al., 1995). The Treiman et al. (1995) analysis provides some evidence that bodies may constitute more frequent constituents than leads, since they observed that the number of different bodies in phonological CVC words was smaller than the number of different leads. Thus the importance of the body units may be related both to their frequency of occurrence in the language and to their contribution to grapho-phonological disambiguation.

Because body units intervene in print-to-sound conversion, the effect of N size could be related to the number of neighbors that share the body with the target. Indeed, a close look at the experimental stimuli reveals that body-neighbors generally constitute the largest subset of orthographic neighborhood. To our knowledge, the partition of the different types of neighbors within the orthographic neighborhood has never been examined in detail. We report below such an analysis for a French corpus of monosyllables.

In addition, a second aim of the present descriptive study was to determine whether there is evidence to support the notion of body/rime units in French. More specifically, we examined two issues: (1) whether the final consonants help determine the vowel pronunciation; (2) whether there are differences in cooccurrence patterns between onset consonants and vowels compared to vowels and coda consonants.

Method and Results

All the analyses reported below used the monosyllabic words from BRULEX, a computerized French lexical database (Content et al., 1990).

Composition of the Neighborhood

In this analysis, we studied the neighborhood composition for all monosyllabic words of four to six letters appearing in the data base ($N = 987$). For each of these words, we computed the number of orthographic and phonographic neighbors, using the whole Brulex data base as reference corpus. The phonographic neighbors were further classified as sharing the same Body, Lead, or Consonant Skeleton.

The pool of words considered in this analysis was restricted to those having at least one phonological initial consonant and one final consonant. Thus phonological CVC, CCVC, CVCC, and CCVCC words were included. Other words (of CV or VC structure) were excluded to ensure that all the words would contribute to the means when averaging the number of neighbors of different types.

Figure 3 depicts the evolution in the mean number of each type of neighbors as a function of orthographic neighborhood size. Two features are worth noting. First, the number of body-neighbors is always larger than the two other types of neighbors. Letter strings with small and large N size essentially differ in the number of body-neighbors, as already suggested by other authors for English (Taft, 1991). Second, the increase in the number of body neighbors as a function of N size is much larger than the augmentation of the number of the other types of neighbors. The latter observation increases the plausibility of the hypothesis that the number of body-neighbors would constitute the relevant factor explaining neighborhood size effects in naming.

Consistency Analysis

Here, we examine whether the additional information provided by the letters following the vowel constrains the pronunciation of the vowel grapheme as it does in English. To better capture the similarities and differences between English and French orthographies with regard to pronunciation consistency, we performed an analysis on a subset of French monosyllabic words similar to the one described by Treiman et al. (1995) for English. The word set consisted of all 772 words whose spoken form consists of a consonant vowel consonant (CVC) sequence. Consistency measures were computed for each graphemic unit[1] (C_1, V, C_2) as well as for units comprising two adjacent graphemes (C_1V and VC_2) on the basis of the whole set of monosyllabic words ($N = 2451$). The *type consistency* measure was the proportion of words containing a given graphemic unit with the same pronunciation as in the target word, relative to the total number of words containing that particular graphemic unit. For instance, the initial consonant of GEL, G, occurs in 70 words. It

[1] As in Treiman's work, orthographic forms were analyzed in terms of graphemes. For instance, U after G or Q (as in GUIDE) was considered part of the C1; the final E, which is not pronounced but ensures that the preceding consonant is pronounced was included in the C2 grapheme. In general, letter groups corresponding to semivowels were treated as consonants. However, there are a few graphemic units in French in which a letter group (OI, OIN) corresponds to a spoken sequence comprising a semivowel and vowel (/**wa**/ and /**wɛ̃**)/, as in FROID, /**frwa**/ or POINT **pwɛ̃**. For these 79 words, the semivowel was considered as part of the vocalic unit.

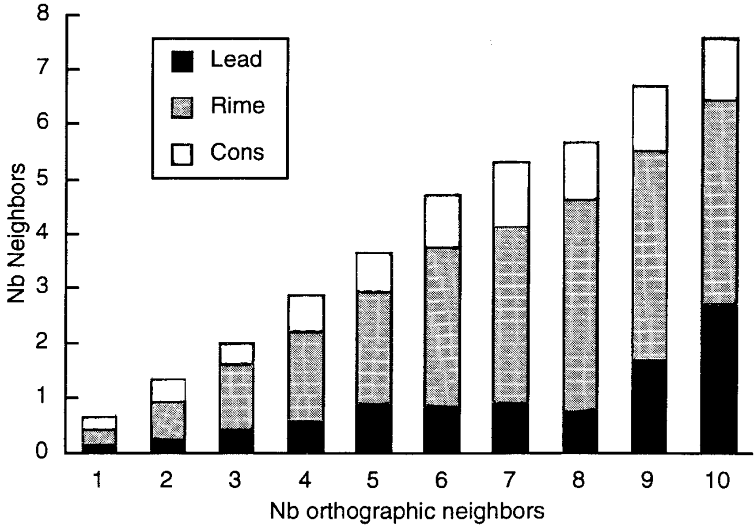

FIG. 3. Composition of the phonographic neighborhood as a function of the number of orthographic neighbors (from 1 to 10 neighbors). The total height of the bar is proportional to the number of phonographic neighbors. Data computed from a sample of 987 monosyllabic words of four to six letters long starting and ending with a consonant (or consonant cluster).

appears in 18 words with the pronunciation /ʒ/ and in 51 words with the pronunciation /**g**/, as in GARE or GOLF, and in one word (GIN) with the pronunciation /**d**ʒ/. Hence, the consistency of the G–/ʒ/ correspondence is .257, and the consistency of the G–/**g**/ correspondence is .729. A similar technique was used to compute *token consistency.* Here, rather than counting the number of words including a particular unit, we summed their frequencies of usage and calculated the summed frequency ratio. Both type and token consistency scores were then averaged for each type of graphemic unit over the 772 CVC words.

The results for French are shown in Table 5 together with the corresponding data from the Treiman et al. study. As can be seen, contrary to English, vowel consistency is very high in French, closely resembling consistencies for initial and final consonants. Moreover, whereas Treiman et al. showed a remarkable increase in consistency when vowel and coda consonants were combined to predict vowel pronunciation, such was not the case in our analysis. Neither did consistency increase more for VC_2 units than for C_1V units as for English.

It should be noted, however, that the pronunciation assigned to one grapheme sometimes depends on the following letters. One case in point which bears on C_1 and V interdependency is the pronunciation of G and C. Another example concerns nasal vowels. The letters N and M after a vowel determine a

TABLE 5

MEAN PERCENTAGES OF PRONUNCIATION CONSISTENCY FOR CVC WORDS IN ENGLISH (FROM TREIMAN *et al.,* 1995) AND IN FRENCH

	By type		By token	
Unit	English	French	English	French
C_1	94	94	96	94
V	62	92	51	91
C_2	92	92	91	91
C_1V	55	95	52	95
VC_2	80	95	77	94

Note. Results are based on 1329 CVC words in English and 772 CVC words in French.

nasal vowel when followed by a consonant, but not when followed by E (compare CAMP, /**kã**/ and DAME, /**dam**/). In the analysis above, N and M were parsed with the vowel in the former, but not in the latter case. If an alternative orthographic parsing procedure is applied, and the letters N and M are considered as part of the C_2 letter group, lead and vowel consistency decrease considerably (78 and 75 for the lead and 75 and 70 for the vowel by type and token, respectively).

Number and Frequency of Lead and Body Units

Treiman et al. (1995) observed that the number of different bodies was smaller than the corresponding number of different leads and that the mean frequency of occurrence of body units (independently of their phonological counterpart) was higher than the mean frequency of lead units. They mentioned the existence of a similar asymmetry between leads and rimes in a parallel analysis of spoken forms and argued that these differences constitute further indications in favor of the body/rime hypothesis.

We counted the number of units of each type (C_1, V, C_2, C_1V, VC_2) independently for the orthographic and the phonological forms. In addition, space occupation ratios were computed to estimate the role of cooccurrence constraints. These ratios correspond to the proportion of *existing* bodies (or leads) relative to *possible* bodies (or leads), that is, the product of the number of V letter groups by the number of C_2 letter groups (or C_1 by V). If no cooccurrence constraints existed, the number of possible and existing combinations should be identical, and the ratio should equal 1. The counts were performed on the whole set of monosyllabic words in BRULEX.

We shall examine first the results of the orthographic counts (Table 6). Contrary to English, the absolute number of different orthographic bodies occurring in French was slightly higher than the number of different leads (883 vs 861 for the whole corpus). Yet, the inverse relationship was observed on the space occupation ratios, suggesting that, as in English, there is more interdependency between constituents of bodies than between constituents of leads. This observation explains why, in the analysis of the composition of the orthographic neighborhood (see Fig. 3), the number of neighbors differing by C_1 (the body-neighbors) was higher than the number of neighbors differing by C_2 (the lead-neighbors). For a given vowel, there are fewer constraints, and hence more different choices, for C_1 than for C_2.

Concerning now the phonological counts, there were much more different leads than rimes (572 vs 469), and the ratio of existing combinations over possible combinations (.252 vs .186) showed a more marked difference. Both the absolute numbers and the ratios indicate the same trend, although it appears less pronounced than in English. Thus, descriptively, the phonological analysis is in agreement with the notion that the rime constitutes a subsyllabic constituent, since it shows that there is more cohesiveness or interdependence between the vowel and final consonant cluster than between the onset and the vowel.[2]

Discussion

At the outset, the differences between English and French orthographies might have led to the undermining of the role of bodies and rime constituents. First, the French orthography is generally considered to provide a much more systematic representation of the phonology than the English writing system. Although exception words do exist, they are less numerous than in English. Thus, if the importance of the body unit is determined by consistency, bodies should play a less important function in French than in English. Second, the predominant syllable structure in French is the open syllable (CV or CCV), whereas in English, open and closed syllables are approximately equally frequent. According to a com-

[2] Similar counts were performed on the subset of 772 words of CVC structure and yielded analogous results.

TABLE 6

NUMBER AND FREQUENCY OF C_1, V, C_2, C_1V, AND VC_2 UNITS

Monosyllables corpus	C_1V	VC_2	C_1	V	C_2
No. of orthographic units	861	883	135	66	220
Space occupation ratio (Orth.)	.096	.061			
No. of phonological units	572	469	108	21	120
Space occupation ratio (Phon.)	.252	.186			
Mean No. of letters	3.16	3.77	2.13	1.92	2.85
Mean No. of phonemes	2.64	3.13	1.92	1.19	2.56
No. graph/No. phon	1.51	1.88	1.25	3.14	1.83

Note. ‘‘C_1’’ and ‘‘C_2’’ columns designate the initial and final consonant or consonant cluster, respectively.

parative token count, the proportion of open syllables amounts to 77% in French and 48% in English (Frauenfelder, Content, Goldman, & Meunier, 1995). This characteristic implies that the frequency of multiphoneme rimes, in both the spoken and the written language, may be lower in French than it is in English. Moreover, a recent study by Taft and Radeau (1995) provides some evidence that the initial syllable constitutes a relevant processing unit in the pronunciation of French written words.

Our analysis of consistency confirmed the high degree of systematicity of the French orthography. However, as we mentioned above, the conclusion dictated by the consistency analysis entirely depends on the way orthographic representations of words are parsed. Hence, it might be premature to conclude that final consonants do not help disambiguate vowel pronunciation in French.

A stronger argument favoring the body/rime specificity can be adduced from the analysis of cooccurrence patterns. Despite the fact that the absolute number of leads was higher than the number of bodies, both the orthographic and the phonological analyses suggest that bodies/rimes constitute more cohesive groups than the leads. This renders the study of sensitivity to subsyllabic constituents in French particularly interesting, since it would permit assessment of the influence of cooccurrence regularities with little contamination of print-to-sound consistency.

Finally, the analyses confirmed that most of the orthographic neighbors are body-neighbors, so that the increase in *N* size was confounded with an increase in the number of body-neighbors. This finding led us to examine experimentally whether body-neighbors would be responsible for the *N* size effect.

EXPERIMENT 3: BODIES AND LEADS

The aim of Experiment 3 was to examine whether body-neighbors play a more important role than other phonographic neighbors in facilitating pronunciation. Thus we contrasted two sets of pseudowords selected so that they differ maximally by the number of body-neighbors. A low neighborhood control set provided a baseline, and both immediate and delayed naming data were collected.

Experiment 3a

Method

Participants. Twenty-one students at the University of Bourgogne participated in the experiment for course credit.

Stimuli and procedure. Three sets of 18 monosyllabic pseudowords were created. Table 7 provides summary statistics for the stimulus sets. Two sets had a similar number of phonographic neighbors, but differed by the proportion of phonographic neighbors that shared the body/rime. Since we tried to keep the number of neighbors sharing the vowel

TABLE 7

CHARACTERISTICS OF THE PSEUDOWORDS USED IN EXPERIMENT 3A (MEAN VALUES)

Characteristics	High body N	Low body N	Controls
No. of neighbors/ON	5.9	5.2	0.3
Freq. of neighbors/ON[a]	1017.8	1160.9	1.8
No. of neighbors/OPN	4.8	3.9	0.1
Freq. of neighbors/OPN[a]	962.9	974.6	0.1
No. of body-neighbors/OPN	3.8	0.7	0.1
Freq. of body-neighbors/OPN[a]	920.0	6.6	0.1
No. of lead-neighbors/OPN	0.2	2.4	0.0
Freq. of lead-neighbors/OPN[a]	1.5	943.5	0.0
No. of vowel-neighbors/OPN	4.0	3.2	0.1
Freq. of vowel-neighbors/OPN[a]	921.4	950.1	0.1
No. of letters	4.7	4.7	4.7
Log bigram frequency	3.2	3.0	3.1

Note. ON, orthographic neighborhood; OPN, phonographic neighborhood.
[a] Mean summed word frequencies (per million) from Imbs (1971).

as constant as possible, this manipulation amounts to contrasting pseudowords that have many lead-neighbors and few body-neighbors with pseudowords having few lead-neighbors and many body-neighbors. For example, three of the four phonographic neighbors of the pseudoword CIVRE (GIVRE, LIVRE, VIVRE) share the rime -IVRE and one shares the lead (CIDRE). Similarly, all of the five phonographic neighbors of the pseudoword DRISE (BRISE, CRISE, FRISE, GRISE, PRISE) share the rime -ISE. By contrast, none of the four phonographic neighbors of the pseudoword FORLE (FORGE, FORCE, FORME, FORTE) and none of the five phonographic neighbors of the pseudoword PLABE (PLACE, PLAGE, PLANE, PLATE, PLEBE) includes the target rime. As in the previous experiments, a set of control pseudowords having few orthographic neighbors was also used. In all other respects, the procedure was identical to Experiment 2. The stimuli are given in Appendix A.

Results and Discussion

Naming latencies falling out of the range between the two cutoff values were excluded from the analyses (0.9 and 2.9%, in immediate and delayed naming, respectively). Latencies corresponding to erroneous triggering of the voice key (3.0 and 3.4%, respectively) or to errors were also omitted from the analyses. The mean naming latencies and the percentages of errors appear in Table 8.

Immediate naming. The effect of Pseudoword Category was significant by subjects ($F_1(2,40) = 25.64$, $p < .001$) but failed to reach significance by items ($F_2(2,51) = 2.56$, $p = .087$). The 28-ms advantage of High-Body pseudowords over the Low-Body pseudowords was reliable by subjects ($F_1(1,40) = 34.52$, $p < .001$) but only marginally signifi-

TABLE 8

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF PSEUDOWORD CATEGORY IN EXPERIMENT 3A

Pseudoword category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
High body N	537	6.3	357	3.4
Low body N	565	7.4	379	1.6
Controls	568	6.3	360	3.2

Note. ON, orthographic neighborhood.

cant by items ($F_2(1,51) = 3.42$, $p = .07$). High-Body pseudowords were 32 ms faster than control pseudowords ($F_1(1,40) = 42.03$, $p < .001$; $F_2(1,51) = 4.22$, $p < .05$). There was no significant difference between Low-Body pseudowords and control pseudowords. The analyses of errors did not show any significant effect.

Delayed naming. As can be seen in Table 8, mean naming latencies differed across pseudoword categories. The effect was reliable by subjects ($F_1(2,40) = 5.82$, $p < .01$) but not by items ($F_2(2,51) = 2.14$, $p = .13$). Although High-Body pseudowords and control pseudowords did not differ significantly, latencies for Low-Body pseudowords were 19 ms longer than those for control pseudowords ($F_1(1,40) = 7.02$, $p < .025$; $F_2(1,51) = 2.72$, $p = .11$) and were 22 ms longer than High-Body pseudowords ($F_1(1,40) = 10.16$, $p < .01$; $F_2(1,51) = 3.64$, $p = .06$). There was no reliable effect of Pseudoword Category in the error analyses ($F_1(2,40) = 2.16$, $p = .13$; $p > .40$ by items).

Differences between immediate and delayed naming latencies. One-way analyses of variance on latency differences between the immediate and the delayed naming task elicited a significant effect of Pseudoword Category ($F_1(2,40) = 7.84$, $p < .01$; $F_2(2,51) = 4.10$, $p < .025$). Both High-Body and Low-Body pseudoword sets yielded smaller differences (180 and 186 ms, respectively) than control pseudowords (208 ms; $F_1(1,40) = 14.08$, $p < .001$; $F_2(1,51) = 7.29$, $p < .01$, for High-Body pseudowords; $F_1(1,40) = 8.86$, $p < .01$; $F_2(1,51) = 4.72$, $p < .05$, for Low-Body pseudowords). The 6-ms difference between High-Body and Low-Body sets did not approach significance ($p > .40$ in both analyses). Thus, when articulatory differences were taken into account, it seems that common bodies do not matter: neighborhood size facilitates performance equally for pseudowords having numerous body-neighbors and for pseudowords having few body-neighbors.

However, because opposite conclusions were reached depending on whether raw immediate naming latencies or difference scores were used as dependent variables, we deemed it appropriate to try and replicate the experiment with a new set of materials.

Experiment 3b

Method

Subjects. Twenty-one undergraduate students at the Free University of Brussels took part in the experiment for partial fulfillment of a course requirement.

Stimuli and Procedure. Three sets of 20 French monosyllabic pseudowords four or five letters long were used (see Appendix A). They were selected on the basis of the same criteria as for Experiment 3a. Descriptive statistics are displayed in Table 9. Six items (one for High-Body, three for Low-Body, and two for control pseudowords) had also been used in Experiment 3a. The procedure was identical to that of Experiment 3a.

Results

The data of two subjects in the immediate naming task were lost due to a malfunction of the data collection program. Hence, the analyses were based on 19 subjects both in immediate naming and delayed naming. Naming latencies falling out of the range between the two cutoff values (0.3 and 5.4% of the trials in immediate and delayed naming, respectively) or corresponding to voice key failures (1.4 and 1.4%, respectively) were excluded from the analyses. One item in each pseudoword category was also removed, one item because it was a conjugated form of a verb (TATE) and the two others (RAIT, LURT) because subjects disagreed about whether or not the final consonant should be pronounced (54 and 46%).

Immediate naming. Pseudoword Category was reliable in the analysis by subjects ($F_1(2,36) = 5.95$, $p < .01$) but not in the analysis by items. Mean naming latencies did not differ significantly between the High-Body and Low-Body pseudoword sets (see Table 10). Low-Body pseudowords were named 20

TABLE 9

CHARACTERISTICS OF THE PSEUDOWORDS USED IN EXPERIMENT 3B (MEAN VALUES)

Characteristics	High body N	Low body N	Controls
No. of neighbors/ON	6.6	6.3	0.2
Freq. of neighbors/ON[a]	1531.2	850.6	2.5
No. of neighbors/OPN	4.6	5.3	0.1
Freq. of neighbors/OPN[a]	1264.5	766.0	1.3
No. of body-neighbors/OPN	3.8	1.4	0.0
Freq. of body-neighbors/OPN[a]	1261.4	20.9	0.0
No. of lead-neighbors/OPN	0.5	2.5	0.0
Freq. of lead-neighbors/OPN[a]	1.2	288.5	0.0
No. of vowel-neighbors/OPN	4.3	3.8	0.0
Freq. of vowel-neighbors/OPN[a]	1262.6	309.5	0.0
No. of letters	4.7	4.7	4.7
Log bigram frequency	3.1	3.2	3.0

Note. ON, orthographic neighborhood; OPN, phonographic neighborhood.
[a] Mean summed word frequencies (per million) from Imbs (1971).

ms faster than control pseudowords. This difference was reliable in the subjects analysis ($F_1(1,36) = 11.62$, $p < .01$) and approached significance in the items analysis ($F_2(1,54) = 3.19$, $p = .08$). The 13-ms difference between pseudowords with many Body neighbors and control pseudowords was significant in the analysis by subjects only ($F_1(1,36) = 4.71$, $p < .05$). There was no reliable effect in the analyses on errors.

Delayed naming. Similar analyses were carried out on the delayed naming latencies and errors. There were no significant effects of Pseudoword Category, either in the analysis on latencies ($p > .20$) or in the analysis on errors ($F_1(2,36) = 1.78$, $p = .18$; $p > .40$ by items).

Differences between immediate and delayed naming latencies. Although a significant effect of Pseudoword Category was observed in the immediate naming task, but not in the delayed naming task, analyses based on the latency differences between immediate and delayed naming failed to show any significant effect, either in the analysis by subjects ($F_1(2,36) = 1.90$, $p = .16$) or in the analysis by items ($p > .30$).

TABLE 10

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF PSEUDOWORD CATEGORY IN EXPERIMENT 3B

Pseudoword category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
High body N	527	6.9	336	1.3
Low body N	520	5.0	341	2.8
Controls	540	7.8	347	2.8

Note. N, neighborhood.

Discussion

The results of Experiment 3b are ambiguous. When immediate naming latencies are analyzed separately from delayed naming performance, it appears that pseudowords having many lead-neighbors are named as fast as pseudowords having many body-neighbors, both sets being faster than controls. On the other hand, when small and nonsignificant differences across stimulus categories in delayed naming are taken into account, then the effect of pseudoword category disappeared. Any-way, whatever analysis is considered, contrary to expectations, no special advantage related to body-neighbors was observed.

However, the evidence in favor of that conclusion suffers from important limitations. It was impossible to manipulate orthogonally the number of lead- and body-neighbors. As shown previously (Fig. 3), most of the phonographic neighbors of a letter string share the body–rime correspondence. Hence, very few words have several lead-neighbors and no body-neighbors. Similarly, several of the pseudowords with many lead-neighbors also had one or even several body-neighbors. Thus, part of the effect for the pseudowords with many lead-neighbors could stem from their body-neighbors. In view of the theoretical importance of this issue, we shall address it through multiple regression analyses.

There is a third type of neighbor that we have not yet examined, namely consonant-neighbors, which share both initial and final consonants, but not the vowel, with the target. In order to conclude that all neighbors are equally important, it should be established that pseudowords having many consonant-neighbors show a similar facilitation effect to the ones having mostly lead- and body-neighbors. Since, by definition, both lead- and body-neighbors must have the vowel in common with the target pseudowords we shall label them vowel-neighbors. Experiment 4 was designed to contrast consonant- and vowel-neighbors.

EXPERIMENT 4: CONSONANT AND VOWEL NEIGHBORS

In Experiment 4, we investigated whether consonant neighbors also give rise to facilitation. Hence, we contrasted two categories of pseudowords with many phonographic neighbors, which differed on the proportion of phonographic neighbors with the same vowel. If the neighborhood effect is contingent on the size of the subset of neighbors that include the same vowel as the target, then only pseudowords with numerous vowel neighbors should show a RT advantage relative to control pseudowords with few or no orthographic neighbors.

Method

Subjects. Twenty-two students at the University of Bourgogne took part in the experiment for course credit. All were native speakers of French.

Stimuli and procedure. Three sets of 19 monosyllabic pseudowords were used (see Appendix A). The two first categories consisted of pseudowords having numerous phonographic neighbors. The first set (High Vowel N) consisted of pseudowords having a large proportion of phonographic neighbors that shared the vowel with the target. For example, the pseudoword NARE has six phonographic neighbors which all include the vowel /**a**/ (GARE, NAGE, LARE, MARE, RARE, TARE). Similarly, all of the phonographic neighbors of the pseudoword VIPE (VICE, VIDE, VILE, PIPE, VITE, VIVE) include the vowel /**i**/. In contrast, the pseudowords of the second set had many phonographic neighbors that did not share the target vowel. For example, the pseudoword RUVE also has six phonographic neighbors, but only three of them include the vowel /**y**/ (RUSE, RUDE, CUVE) and the remaining ones differ by the vowel (RIVE, REVE, RAVE). Similarly, the pseudoword MUME has five phonographic neighbors (MIME, MULE, MUSE, MEME, MOME), but only two of them share the target vowel. The third category included pseudowords with few orthographic neighbors. Summary statistics appear in Table 11. The list of stimuli started with one warm-up trial and was preceded by 20 practice trials. Other details of the procedure were as previously described.

Results

Naming latencies corresponding to erroneous triggering of the voice key (2.4 and 4.5% of the trials in immediate and delayed naming, respectively), or falling out of the range between the two cutoff values (0.3 and 4.9%,

TABLE 11

CHARACTERISTICS OF THE PSEUDOWORDS USED IN EXPERIMENT 4 (MEAN VALUES)

Characteristics	High vowel N	Low vowel N	Controls
No. of neighbors/ON	5.8	5.0	0.3
Freq. of neighbors/ON[a]	814.8	1200.3	7.6
No. of neighbors/OPN	5.2	3.9	0.2
Freq. of neighbors/OPN[a]	808.8	858.1	6.3
No. of body-neighbors/OPN	3.2	1.3	0.0
Freq. of body-neighbors/OPN[a]	624.4	28.5	0.0
No. of lead-neighbors/OPN	1.8	0.9	0.1
Freq. of lead-neighbors/OPN[a]	178.8	13.4	1.3
No. of vowel-neighbors/OPN	5.0	2.3	0.1
Freq. of vowel-neighbors/OPN[a]	803.2	41.9	1.3
No. of letters	4.6	4.6	4.6
Log bigram frequency	3.2	3.0	3.1

Note. ON, orthographic neighborhood; OPN, phonographic neighborhood.
[a] Mean summed word frequencies (per million) from Imbs (1971).

respectively) were left out from the analysis (see Table 12).

Immediate naming. There was a reliable effect of Pseudoword Category ($F_1(2,42) = 34.45$, $p < .001$; $F_2(2,54) = 6.91$, $p < .01$). The High-Vowel pseudowords were pronounced 43 ms faster than the controls ($F_1(1,42) = 68.27$, $p < .001$; $F_2(1,54) = 13.66$, $p < .001$) and were pronounced 25 ms faster than the Low-Vowel pseudowords ($F_1(1,42) = 23.22$, $p < .001$; $F_2(1,54) = 4.75$, $p < .05$). The 18-ms difference between the control pseudowords and the Low-Vowel pseudowords reached significance in the analysis by-subject only ($F_1(1,42) = 11.85$, $p < .01$; $F_2(1,54) = 2.30$, $p = .14$). No significant effect was observed in the error analyses.

No significant effect was observed in analyses of delayed naming data.

TABLE 12

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF PSEUDOWORD CATEGORY IN EXPERIMENT 4

Pseudoword category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
High vowel N	529	6.7	349	1.2
Low vowel N	554	6.5	355	1.4
Controls	572	7.7	355	3.3

Note. N, neighborhood.

Discussion

The results reveal a clear effect of common-vowel pseudowords. The small effect observed for the pseudowords having many consonant-neighbors, which was significant only in the analyses by subjects, is likely to be due to residual differences in the number of vowel neighbors. Indeed, as indicated in Table 11, the latter set of pseudowords had an average of 2.3 vowel neighbors. Hence, one conclusion that stems from the present experiment is that neighbors having only the consonantal skeleton in common with the target do not influence phonological conversion.

The absence of a significant facilitation effect for the pseudowords with consonant-neighbors indicates that all neighbors are not equally important. Two subsets of pseudowords have been found to give rise to facilitation effects, those having many lead-neighbors and those having many body-neighbors. One possible interpretation of these findings is that the neighborhood effect is determined by the existence of many neighbors having one of these particular subsyllabic units in common with the target. However, a second possibility is that what determines N size effects is not the occurrence of large size units (lead or body) in the phonographic neighbors, but the presence of an identical vowel in the pseudowords and their neighbors. Because it is impossible to completely disentangle these factors experimentally, this issue will be taken up through regression analyses in which we examine whether the number of lead- and body-neighbors independently contribute to variation in naming times, or whether the total number of vowel-neighbors is the best predictor of performance.

GLOBAL ANALYSES

Before we set out to summarize our findings and discuss their interpretation and implications, we shall present supplementary analyses combining the data of Experiments 2 to 4. Such global analyses are motivated by the relatively high degree of correlation of the factors that were manipulated. Despite all efforts toward isolating independent variables through experimental manipulations, the tables describing stimulus characteristics show that it was generally impossible to dissociate variables completely. Furthermore, the use of regression techniques offered the opportunity to verify the validity of our conclusions on a larger data set, collected on similar samples of subjects and with a homogeneous methodology.

We performed regression analyses to investigate the role of the different predictor variables across all experiments in which the nature of phonographic neighborhood was manipulated (Experiments 2 to 4). Mean immediate and delayed naming times by items were used. Because mean latencies for different items were based on different groups of subjects, we first conducted analyses of variance on latencies for the control pseudowords to examine whether the different groups differed in overall speed. The results showed no indication of group differences, either for immediate naming or for delayed naming ($p = .47$ and $p = .32$, respectively). None of the pairwise comparisons between groups reached significance.

Some letter strings appeared in more than one experiment. In that case, the RT used in the regression analyses was chosen at random, providing a total set of 183 data points. Furthermore, to eliminate the variability due to articulatory execution, we computed a simple regression on immediate naming latency, using mean delayed naming time as predictor. This variable accounted for 41.2% of the variance, and the residual corrected naming time (CNT) was used as dependent variable in subsequent analyses.

We first examined the role of phonographic neighborhood. A preliminary analysis indicates a significant simple correlation between CNT and the number of phonographic neighbors ($r = -.346$, $t(181) = -4.97$, $p < .0001$). A multiple regression using the number of orthographic neighbors, the number of phonological neighbors, the number of phonographic neighbors, and the mean log bigram frequency and the pseudoword length (number of letters) as predictors confirmed the hypothesis: the number of phonographic neighbors was the only significant unique predictor of CNT (see Table 13; partial correlation $-.160$, $t(177) = -2.30$, $p < .025$). None of the other predictors had a significant contribution. This result reinforces the conclusion of Experiments 1 and 2 that the neighborhood facilitation effect is controlled by the size of the set of phonographic neighbors.

A second analysis examined the role of grapho-phonological consistency. We computed consistency scores for all units (C_1, V, C_2, C_1V, VC_2) of each stimulus, based on the statistics described previously. Both Type and Token consistency measures were used, thus providing a total of 10 predictors. There was little evidence that consistency played any role. Only 2 among the 10 consistency vari-

TABLE 13

RAW CORRELATIONS BETWEEN CNT AND THE FIVE INDEPENDENT PREDICTOR VARIABLES CONSIDERED

Variables	Length	Bigrams	NOP	NO	NP
CNT	.1035	−.0452	−.3464**	−.3002**	−.2867**
Length		.3365**	−.0783	−.1952**	−.3416**
Bigrams			.0815	.0568	.1360
NOP				.8528**	.4502**
NO					.4703**
Part Cor	.0352	−.0143	−.1600*	.0247	−.1218
Beta	.0421	−.0159	−.3156	.0491	−.1512

Note. NOP, number of phonographic neighbors; NO, number of orthographic neighbors; NP, number of phonological neighbors. The two lower lines display the partial correlations of each predictor with CNT, and the standardized regression coefficients.

* $p < .05$, two-tailed test.

** $p < .01$, two-tailed test.

ables showed significant simple correlations with CNT (C_1Type and C_1Token: $r = -.281$, $t(173) = -3.85$, $p < .001$; $r = -.205$, $t(173) = -2.75$, $p < .01$, respectively). Separate multiple regressions, pitting each of these variables against the number of phonographic neighbors, indicated that only C_1Type and C_1Token significantly increased the proportion of explained variance[3] (partial correlations: $-.264$, $t(172) = -3.81$, $p < .001$; $-.187$, $t(172) = -2.64$, $p < .01$, respectively). Together, C_1Type and the number of phonographic neighbors accounted for 17.5% of the variance. The corresponding value for C_1Token and the number of phonographic neighbors was 14.0%. Finally, the influence of the number of phonographic neighbors was still significant ($r = -.34$, $t(126) = -4.01$, $p < .001$) if the data set was restricted to the items for which Body consistency was perfect.

The main conclusion of this first set of analyses is that the facilitatory effect of N size cannot be explained by consistency. In a large-scale experiment, Treiman et al. (1995) found that naming performance was affected by the consistency of both C_1 and the body (VC_2). Although we observed a similar effect of C_1 consistency on naming pseudowords, there was no effect of body consistency. The limited influence of consistency is perhaps less surprising if one remembers that most of the stimuli were highly consistent. In terms of the Body/Rime unit, only seven pseudowords had a consistency score lower than 95. Hence, the detrimental effect caused by the inconsistency of body−rime correspondences should be more easily observable for English than for French.

The last set of analyses was aimed at clarifying the influence of the different subsets of phonographic neighbors. Table 14 displays the raw correlations between CNT and the predictors that were examined. The total number of phonographic neighbors (NOP), the number of vowel-neighbors (NV) and the number of body-neighbors (NB) are highly intercorrelated and show similar correlations with CNT, each accounting for 11 to 12% of the variance. This is not surprising, given that our lexical analyses (see Fig. 3) showed that most phonographic neighbors share either the body or the lead with targets. What appears more surprising is that, contrary to the conclu-

[3] Separate analyses were more appropriate because all consistency scores could not be defined for all items, either because a given unit did not exist in the lexical data base or because (for the scores based on token counts) the relevant words usage frequency was not available. All consistency scores were available for a subset of 111 of the 183 items.

TABLE 14

RAW CORRELATIONS BETWEEN CNT AND NEIGHBORHOOD PARAMETERS

Variables	NOP	NCons	NVowel	NLead	NBody
CNT	−.3464**	−.1571*	−.3347**	−.1058	−.3328**
NOP		.4793**	.9511**	.5558**	.7756**
NCons			.2142**	.2017**	.1236
NVowel				.5580**	.8329**
NLead					.0056

Note. NOP, number of phonographic neighbors; NCons, number of consonant-neighbors; NVowel, number of vowel-neighbors; NLead, number of lead-neighbors; NBody, number of body-neighbors.

* $p < .05$, two-tailed test.

** $p < .01$, two-tailed test.

sions of Experiment 3, the number of lead-neighbors (NL) does not correlate significantly with CNT.

Despite the high intercorrelations between NOP, NB, and NV, we attempted to determine which of these three variables provides the best account of the effect, by using the following strategy. Recall that NOP is the sum of NV and NC, the number of Consonant neighbors. If the facilitation effect is controlled by the total number of phonographic neighbors, one would expect NC to increase the proportion of explained variance after NV has been forced into the equation. In fact, after NV was entered, NC made no significant unique contribution (partial correlation .09, $t = 1.25$, $p > .20$). Conversely, after NC was forced into the equation, accounting for 2.5% of the variance, NV still had a unique contribution, accounting for a further 9.6% of the variance (partial correlation −.312, $t = 4.41$, $p < .001$).

Similarly, NV is the sum of NB and NL. If the total number of vowel-neighbors, NV, determines the effect, one would expect both NB and NL to contribute. However, NL had no significant contribution after NB was entered in the equation (partial correlation = −.106, $t = 1.49$, $p > .10$), whereas NB contributed a highly significant portion of explained variance after NL was forced into the equation (partial correlation = −. 332, $t = 4.76$, $p < .001$). In sum, when partitioning the set of vowel-neighbors into those sharing the lead and those sharing the body, the analysis suggests that the size of the body neighborhood is much more related to performance than the size of the lead neighborhood, and that the size of the body neighborhood is as closely related to performance as the size of the vowel neighborhood.

Finally, the conclusion of the regression analyses stands in contradiction with the outcome of the analysis of variance on Experiment 3, which indicated similar facilitation for lead and body neighborhoods. However, as can be seen in Tables 7 and 9, the number of body-neighbors for the low body-neighborhood stimulus sets was still higher than the number of body-neighbors in the control pseudowords, and this difference may account for the effect which we initially attributed to the number of lead neighbors.

This interpretation implies that the existence of very few body-neighbors should be sufficient to cause facilitation. We thus reanalyzed the data of each experiment according to two dichotomous criteria, the existence of more than one body-neighbor and the existence of more than one lead-neighbor. In each case, the existence of body-neighbors facilitates pseudoword naming, whereas the existence of lead-neighbors showed no significant effect at all. Indeed, two-way analyses of variance indicated that the number of body-neighbors was the only significant factor in each case ($F(1,50) = 4.31$, $p < .05$; $F(1,50) = 4.95$,

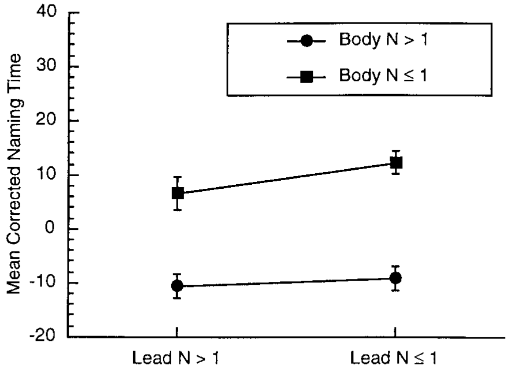

FIG. 4. Mean Corrected Naming time as a function of the size of body-neighborhood (Body N) and lead-neighbors (Lead N). Data are pooled from Experiments 2, 3, and 4. The vertical bars represent one standard error.

$p < .05$; $F(1,53) = 3.05$, $p = .086$; $F(1,53) = 9.32$, $p < .005$, respectively, for Experiments 2, 3a, 3b, and 4). Neither the existence of lead-neighbors nor the interaction reached significance. Figure 4 displays the mean results pooled from the four experiments.

EXPERIMENT 5: THE BODY NEIGHBORHOOD OF WORDS

One potential limitation of the previous experiments stems from the use of pseudowords. One could wonder whether the effects depend on the nonlexical nature of the letter strings. Neighborhood facilitation effects have rarely been observed with high frequency words. For low-frequency items, however, neighborhood size effects have been demonstrated in English (Andrews, 1989, 1992, Sears et al., 1995) as well as in French (Peereman & Content, 1995). In the latter study, similar effects were obtained for low-frequency items and for pseudowords.

While these results suggest that similar patterns should obtain for pseudowords and low-frequency words, no direct evidence on the influence of the body neighborhood size is available. In addition, given the divergence between the conclusion of factorial manipulations and regression analyses, it seemed appropriate to test directly the effect of body neighborhood on real words.

In the last experiment, we examined whether an effect of the number of body-neighbors could be observed for real words. Two sets of words with many phonographic neighbors were compared, which differed specifically on the proportion of body-neighbors. If the neighborhood effect is contingent on the size of the body neighborhood, then only words with numerous body-neighbors should show a RT advantage when compared to control words with few or no orthographic neighbors.

Method

Subjects. Fifty-three students at the University of Bruxelles took part in the experiment for course credit. All were fluent speakers of French.

Stimuli and procedure. Three sets of 25 monosyllabic words four to six letters long were used (see Appendix A). The two first sets consisted of words having numerous phonographic neighbors. The words in the first set (High Body N (High BN)) had a high proportion of phonographic neighbors that shared the body with the target, whereas those in the second set (Low Body N (Low BN)) had few phonographic neighbors sharing the body. The third category included words with few or no orthographic neighbors. High BN words were selected among those words having at least five body-neighbors. Based on the results of the reanalysis of the pseudoword data, Low BN words were selected from among words having at least five orthographic neighbors and less than two body neighbors. Low N Controls had less than two orthographic neighbors. Summary statistics appear in Table 15.

The list of stimuli started with two warm-up trials and was preceded by 20 practice trials. Other details of the procedure were identical to those of Experiments 1 and 2. All subjects were first tested on the immediate naming condition and then, after a short break, on the delayed naming condition.

TABLE 15

CHARACTERISTICS OF THE WORDS USED IN EXPERIMENT 5 (MEAN VALUES)

Characteristics	High body N	Low body N	Controls
No. of neighbors/ON	7.3	6.4	0.3
Freq. of neighbors/ON[a]	517.5	583.3	7.8
No. of neighbors/OPN	6.6	3.3	0.3
Freq. of neighbors/OPN[a]	494.5	100.3	7.5
No. of body neighbors/OPN	5.4	0.6	0.1
Freq. of body neighbors/OPN[a]	470.5	37.7	0.1
No. of lead neighbors/OPN	0.4	1.6	0.1
Freq. of lead neighbors/OPN[a]	5.9	34.8	6.4
No. of vowel neighbors/OPN	5.8	2.2	0.2
Freq. of vowel neighbors/OPN[a]	476.4	72.5	6.6
No. of letters	4.9	4.7	5.5
Log bigram frequency[b]	3.1	3.0	3.0
Word frequency[c]	8.7	11.0	4.1

Note. ON, orthographic neighborhood; OPN, phonographic neighborhood.
[a] Mean summed word frequencies (per million) from Imbs (1971).
[b] Bigram frequencies were taken from Content and Radeau (1988).
[c] Mean word frequencies (per million) from Imbs (1971).

Results

Due to random microphone malfunctionings, approximately 10% of the response times were not recorded (9.3 and 8.9% in the immediate and delayed condition, respectively). Application of latency cutoff values and invalid triggerings of the voice key led to the rejection of 2.4% of the trials in the immediate naming task and 3.1% in the delayed naming task. Because the three sets were not perfectly matched for word frequency, it was entered as a covariant in the item analyses. The mean naming latencies and the percentages of errors appear in Table 16.

Immediate naming. There was a reliable effect of Stimulus Category ($F_1(2,104) = 45.02$, $p < .0001$; $F_2(2,71) = 3.15$, $p < .05$). The High Body *N* words were pronounced 25 ms faster than the control words ($F_1(1,104) = 86.90$, $p < .0001$; $F_2(1,71) = 5.92$, $p < .025$). The difference between the Low Body N words and controls was marginally significant ($F_1(1,104) = 38.36$, $p < .0001$; $F_2(1,71) = 3.23$, $p = .08$). The 8-ms difference between High Body N and Low Body N words was significant by Subjects only ($F_1(1,104) = 9.79$, $p < .005$; $F_2(1,71) < 1$). No significant effect was obtained in the analyses of error rates ($F_1(2,104) = 2.61$, $p = .08$; $F_2(2,72) < 1$).

Delayed naming. There was a reliable effect of Stimulus Category in the analysis by Subjects ($F_1(2,104) = 5.64$, $p < .005$; $F_2(2,71) = 0.99$). Local comparisons showed that the Low Body N words were pronounced faster than both the High Body N words ($F_1(1,104)$

TABLE 16

MEAN NAMING LATENCIES (IN MS) AND PERCENTAGE OF ERRORS AS A FUNCTION OF STIMULUS CATEGORY IN EXPERIMENT 5

Stimulus category	Immediate naming		Delayed naming	
	Latencies	Errors	Latencies	Errors
High body N	553	3.6	435	0.5
Low body N	561	3.6	424	0.8
Controls	578	2.4	433	0.2

Note. N, neighborhood.

$= 6.51$, $p = .01$) and the controls ($F_1(1,104) = 10.0$, $p = .002$), which did not differ from each other. No significant effect was obtained in the analyses of error rates ($F_1(2,104) = 2.66$, $p = .08$; $F_2(2,72) = 2.2$, $p > .10$).

Differences. The analyses on difference scores showed a reliable effect of Stimulus Category ($F_1(2,104) = 14.94$, $p < .0001$; $F_2(2,71) = 5.80$, $p < .005$). The High Body N words were read 23 ms faster than the control words ($F_1(1,104) = 27.5$, $p < .0001$; $F_2(1,71) = 11.2$, $p = .001$). The 17-ms difference between High Body N and Low Body N words was also reliable ($F_1(1,104) = 15.68$, $p < .0001$; $F_2(1,71) = 4.59$, $p < .05$). The 6-ms difference between the Low Body N words and controls did not reach significance ($F_1(1,104) = 1.65$, $p > .20$; $F_2(1,71) = 1.45$, $p > .20$).

Discussion

Immediate naming latencies showed only a nonsignificant trend for High Body N words to be pronounced faster than Low Body N words. Delayed naming latencies showed an advantage for the Low Body N words, which was significant only in the Subject analysis. This suggests that the difference is not systematically related to stimulus categories, but rather that it is due to spurious variations among items in articulatory execution or in word onset energy. When these factors are eliminated by taking the immediate–delayed naming latency differences as dependent variable, the results unequivocally support our previous conclusions: High Body N words were processed significantly faster than those in the two other stimulus categories, which did not differ from each other. In sum, the results provide some support for the claim that word naming is sensitive to the number of neighbors that share the rime.

GENERAL DISCUSSION

The purpose of the present study was to investigate the locus of the neighborhood effect in the naming task. Experiment 1 showed that the size of the phonological neighborhood did not affect RTs when pseudowords were matched on orthographic neighborhood size. This result is in accordance with the absence of a phonological rime frequency effect in word naming reported by Brown and Watson (1994) in English.

Experiment 2 examined the influence of the set of words that are simultaneously orthographic and phonological neighbors, which we have called the *phonographic* neighborhood. The results showed that the facilitatory effect was controlled by the phonographic neighborhood. A regression analysis on the pseudowords employed in Experiments 2 to 4 confirmed this outcome. Among the three estimates of neighborhood density, the number of phonographic neighbors was the only significant unique predictor of naming times. These findings demonstrate that the size of the orthographic neighborhood is a necessary specification, but not a condition sufficient to warrant the occurrence of a facilitation effect. Purely orthographic variables, such as average bigram frequency, or the number of orthographic neighbors did not determine systematic variations in naming performance.

An influential account of the N size effect holds that orthographic neighbors facilitate orthographic encoding (Andrews, 1989, 1992). Previous studies showing that neighborhood size facilitates naming as well as lexical decision performance (e.g., Andrews, 1989, 1992; Forster & Shen, 1996; McCann & Besner, 1987; Peereman & Content, 1995; Sears et al., 1995) did not distinguish orthographic neighbors in terms of their phonological properties. The finding that only the phonographic neighbors play a role in naming is at odds with the orthographic encoding hypothesis. The phonographic effect on pseudowords suggests that lexical neighbors facilitate the generation of a phonological representation through the activation of convergent lexical codes. Whether, in addition, neighborhood accelerates the lexical retrieval process for real words remains an open issue, although we suspect that such a contribution might be counterbalanced by lexical competition effects (Jacobs & Grainger, 1992).

It is conceivable that some part of the neighborhood effect observed in the lexical decision task is also due to phonological activation. In recent years, many authors have shown that phonological information contributes to word identification (Grainger & Ferrand, 1994; Perfetti & Bell, 1991; Perfetti, Bell, & Delaney, 1988; Van Orden, 1987; Verstaen, Humphreys, Olson & d'Ydewalle, 1995). If so, one would expect that the predominant role of phonographic neighbors would also be observed in a lexical decision experiment.

The subsequent experiments examined whether all phonographic neighbors have the same importance in determining naming performance. In Experiments 3 and 4 we tried to contrast the contribution of different subsets of phonographic neighbors which shared either the lead, the body, or the consonantal skeleton with the targets. In Experiment 4, a large facilitatory effect occurred for the stimuli having numerous lead- and body-neighbors, but not for those having predominantly consonant neighbors. Experiment 3 produced mixed results and failed to differentiate between lead- and body-neighbors.

Given the difficulty of factorially manipulating the number of lead and body neighbors, we reexamined the pseudoword data by means of multiple regression techniques. The analysis demonstrated that the number of body-neighbors was the critical factor. The facilitation effect attributed to the number of lead-neighbors was actually better explained by the existence, for these sets of pseudowords, of a small number of body-neighbors. Based on these results, we sorted the stimuli from each experiment according to the number of lead- and body-neighbors, but using more stringent criteria than we had initially adopted. Analyses of variance confirmed the influence of body-neighbors and failed to show any significant contribution of lead-neighbors. Finally, Experiment 5 contrasted words varying by the size of the body neighborhood and supported the conclusion in showing a significant benefit specifically associated with the size of the body neighborhood.

The results can be accounted for by several information processing theories of word recognition and naming. Multiple-route theories assume that body–rime correspondences play an important role in print-to-sound conversion (Norris, 1994; Patterson & Morton, 1985; Shallice & McCarthy, 1985; Taft, 1991). In this framework, body-neighbors would boost the activation of the body and rime units. A letter string having only lead-neighbors would not benefit from a similar advantage since the lead does not correspond to a particular unit in the system. By contrast, it seems unlikely that the Dual Route Cascaded (DRC) model (Coltheart et al., 1993) would be capable of simulating the body neighborhood advantage. The model incorporates excitatory links between the lexical system and the phoneme system. This pathway could perhaps account for the contribution of phonographic neighbors to naming, since, particularly for consistent items, these entries would help activating convergent phonological codes. However, there is no provision in the DRC model to account for differential influences of lead- and body-neighbors. Indeed, the contribution of orthographic neighbors to the activation of the appropriate phonological codes would only depend on the number of phonemes shared with the target letter string.

In the framework of parallel distributed models of print-to-speech transcoding (Seidenberg & McClelland, 1989, Plaut, McClelland, Seidenberg, & Patterson, 1996; Van Orden, Pennington, & Stone, 1990), the neighborhood facilitation effect has been attributed to the strength of connections between orthographic and phonological units. Letter strings from dense neighborhoods are constituted of more frequent letter and phoneme groups. Because the connection strengths should be sensitive to the frequency of cooccurrence of letter groups and phonemic correspondence, the facilitatory effect should be more pronounced when the orthographic neighbors are also phonologically similar to the target. As a consequence, we believe that the restriction of neighborhood facilitation to the phonographic subset emerges as a prediction of these models, although it has not been explicitly spelled out.

In these models, the stipulation of preexisting linguistic constituents is generally avoided to minimize a priori assumptions in the modeling endeavor. However, the nature of learning rules and the training experience yield sensitivity to particular syllable constituents. Indeed, Seidenberg and McClelland (1989) indicated that their model captured the particular importance of word bodies. Similarly, Plaut et al. (1996) noticed that their model picked up the interdependency between vowels and codas, although in that model sensitivity to the syllabic structure might be partly induced by coding decisions. Two kinds of statistical regularities in the training corpus might contribute to determine a particular sensitivity to the body. First, as recently demonstrated by Treiman et al. (1995; see also Stanback, 1992), coda consonants strongly constrain vowel pronunciation. Second, vowel and coda consonants tend to cooccur more than vowel and onset consonants. As shown in the descriptive analyses of our lexical corpus, both at the orthographic and at the phonological level, there is more cohesiveness between rime constituents than between lead constituents. Similar observations have been recently reported by Kessler and Treiman (in press) for phonological forms in English. While the first kind of regularity holds for English much more than for French, due to the high degree of consistency of the French orthography, the cooccurrence asymmetry might enhance sensitivity to bodies and rimes. Interestingly, the Plaut et al. analyses suggest that some interdependency between vowel and coda units develops even for regular and consistent words.

Descriptively, the main conclusion of the study is that not all orthographic neighbors are equally influential. By introducing a refined classification of orthographic neighbors as a function of their phonological properties, we were able to demonstrate that only a subset of neighbors, namely the words that share the body and the corresponding rime, are clearly relevant to naming. Despite the fact that most of the collected evidence stems from pseudoword naming, the final experiment suggests that the conclusion extends to real words. It remains to be seen whether similar phenomena would arise in languages other than French.

The present findings add further weight to the substantial collection of linguistic and psycholinguistic arguments demonstrating the validity of the description of syllables into onset and rime constituents (see, e.g., Treiman (1992) for a review). To our knowledge, this study provides the first psycholinguistic demonstration of the importance of the body/rime unit in the French language. This is particularly interesting given the differences between French and English. It is generally admitted that the French orthography is highly consistent, as far as grapho-phonological transcoding is concerned, and this was confirmed by our statistical analyses. The present evidence thus suggests that the special status of orthographic bodies is not necessarily related to spelling-to-sound consistency. Although consistency may be relevant, at least in some orthographies, it is not a condition necessary for the emergence of sensitivity to the body. Another factor that might contribute is the higher cohesiveness of adjacent graphemes (as well as phonemes) when they are constituents of a body (or rime) than when they are parts of a lead. Indeed, both from an orthographic and a phonological standpoint, the analysis of French monosyllabic words confirmed that mutual constraints between the vowel and the subsequent consonants are stronger than between the initial consonants and the vowel.

APPENDIX A

Experiment 1

ON+PN+: banne, boute, catte, corte, dage, doupe, faine, falle, gace, laire, lette, mide, monne, rure, sare, sonte, taute, toche, velle, voule.

ON−PN+: baime, beire, bène, ceife, ceppe, chice, dret, fiffe, fraut, furre, goul, junne, krair, lorre, naile, nife, paice, puppe, quice, trée.

Controls: bocre, clade, covre, derpe, drir, drour, flide, frude, gume, lapte, loibe, melte, muic, rilt, slode, stue, teude, tunde, vlade, vrane, bline, blir, brouc, caple, ceuge, chuil, crode, dult, frafe, froun, frupe, gufe, ilge, junde, lumbe, naude, plide, plope, stume, ulte.

Experiment 2

OPN+: vorte, baire, jou, nitre, rie, mour, nur, mel, feste, doupe, dil, delle, sare, dige, vipe, nire, dage, boute.

OPN−: aigne, foite, mox, oile, cau, lutre, tos, poite, vigle, cutte, jot, flire, souf, juie, oure, gas, blef, molt.

Controls: virce, glir, eude, vup, froul, aup, sle, éce, tuife, covre, bluir, bro, boipe, flade, nume, ubre, vume, vril.

Experiment 3a

High Body N: beste, civre, conde, derre, doir, drise, felle, gable, honne, jort, lette, nite, noute, puite, rère, bigne, toin, vorte.

Low Body N: chore, dife, doude, faige, forle, jeule, jouf, lipre, monle, plabe, polte, porge, prine, pège, ribe, sorpe, vige, vigre.

Controls: blare, disme, dron, dupre, flade, froge, froul, girt, huble, laufe, mugne, nème, podre, raur, sirde, tudre, virce, vril.

Experiment 3b

High Body N: cutte, deste, feste, figne, gire, jote, juite, londe, lorte, luste, noute, piel, rait, rour, sotre, suir, tause, teine, toir, vrise.

Low Body N: chore, coume, doube, farme, fate, forne, joune, lavre, lipre, lure, monce, pirte, plame, pège, rise, rore, sorde, tate, vatre, vipe.

Controls: caive, ceuge, chide, chée, doime, froul, fruce, jonre, laufe, leuce, lurt, lète, moun, nirte, prafe, prel, reue, tinse, voube, vume.

Experiment 4

High Vowel N: bère, dace, daire, deste, divre, doune, forpe, lide, lite, londe, lonne, nare, norte, noute, puite, relle, rice, sorde, vipe.

Low Vowel N: boune, jope, chase, compe, coume, farme, flite, juie, mande, moude, mume, prin, purte, raute, ruve, vate, vatre, viste, vran.

Controls: blir, ceuge, chont, covre, doime, dron, flade, froge, fruce, jirt, laufe, lète, moun, nirte, ploit, raime, reue, voube, vume.

Experiment 5

High Body N: coche, datte, derme, fendre, gage, gendre, germe, haie, jatte, joule, lasse, latte, leste, mage, nain, natte, peste, rente, roche, rouer, sente, souche, tente, tire, zeste.

Low Body N: bille, bourre, brin, butte, cane, carpe, comte, dune, fosse, frire, halle, hotte, houle, huer, lange, luge, mine, moite, pince, pipe, plaie, rive, saute, troc, verse.

Controls: blâme, buffle, cancre, chiot, crampe, crêpe, dinde, douane, filtre, frein, gifle, guêpe, huître, juive, lèpre, mixte, plâtre, poivre, sieste, sobre, trèfle, triche, truite, valse, zèbre.

REFERENCES

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition,* **10,** 565–575.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **15,** 802–814.

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

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. (1984). Acquisition and use of spelling–sound correspondences in reading. *Journal of Experimental Child Psychology,* **38,** 114–133.

Bowey, J. A., & Hansen, J. (1994). The Development of Orthographic Rimes as Units of Word Recognition. *Journal of Experimental Child Psychology,* **58,** 465–488.

Brown, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language,* **26,** 1–23.

Brown, G. D. A., & Watson, F. L. (1994). Spelling-to-sound effects in single-word reading. *British Journal of Psychology,* **85,** 181–202.

Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology,* **26,** 429–438.

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

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance* (Vol. VI, pp. 535–555). London: Academic Press.

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

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.

Content, A., & Radeau, M. (1988). Données statistiques sur la structure orthographique du français. *Cahiers de Psychologie Cognitive,* **8,** 399–404.

Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 127–146). London: Erlbaum.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **22,** 696–713.

Frauenfelder, U. H., Content, A., Goldman, J.-P., & Meunier, C. (1995). *Comparative sublexical statistics: The processing units debate.* [Paper presented at the 8th annual CUNY Conference, Tucson, AZ, March 1995]

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

Glushko, R. J. (1981). Principles for pronouncing print: the psychology of phonography. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 61–83). Hillsdale, NJ: Erlbaum.

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

Grainger, J., & Ferrand, L. (1994). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language.*

Imbs, P. (1971). *Etudes statistiques sur le vocabulaire français. Dictionnaire des fréquences. Vocabulaire littéraire des XIXe et XXe siècles.* Centre de Recherche pour un trésor de la langue française (CNRS), Nancy, Paris: Librairie Marcel Didier.

Jacobs, A. M., & Grainger, J. (1992). Testing a semistochastic variant of the Interactive Activation Model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 1174–1188.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language,* **22,** 687–715.

Kay, J., & Bishop, D. (1987). Anatomical differences between nose, palm and foot, or, the body in question: Further discussion of the processes of sub-lexical spelling-sound translation. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 449–469). London: Erlbaum.

Kay, J., & Marcel, T. (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology,* **33A,** 397–413.

Kessler, B., & Treiman, R. (in press). Syllable structure and phoneme distribution. *Journal of Memory and Language.*

Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime and the onset of reading. *Journal of Experimental Child Psychology,* **48,** 224–245.

Laxon, V., Masterson, J., & Coltheart, V. (1991). Some bodies are easier to read: The effect of consistency and regularity on children's reading. *Quarterly Journal of Experimental Psychology,* **43A,** 793–824.

Laxon, V., Masterson, J., & Moran, R. (1994). Are Children's Representations of Words Distributed? Effects of Orthographic Neighbourhood Size, Consistency and Regularity of Naming. *Language and Cognitive Processes,* **9,** 1–27.

Laxon, V., Masterson, J., Pool, M., & Keating, C. (1992). Nonword naming: Further exploration of the pseudohomophone effect in terms of orthographic neighborhood size, graphemic changes, spelling–sound consistency, and reader accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **18,** 730–748.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance,* **13,** 14–24.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychological Review,* **88,** 375–405.

Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance,* **20,** 1212–1232.

Paap, K. R., McDonald, J. E., Schvaneveldt, R. W., & Noel, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 221–243). London: Erlbaum.

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia, Neuropsychological and cognitive studies of phonological reading* (pp. 335–359). Hillsdale, NJ: Erlbaum.

Peereman, R. (1995). Naming regular and exception words: Further examination of the effect of phonological dissension among lexical neighbours. *European Journal of Cognitive Psychology,* **7,** 307–330.

Peereman, R., & Content, A. (1995). The neighborhood size effect in naming: Lexical activation or sublexical correspondences. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **21,** 409–421.

Perfetti, C. A., & Bell, L. (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.

Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language,* **27,** 59–70.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review,* **103,** 56–115.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory and Cognition,* **13,** 90–99.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* **21,** 876–900.

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

Seidenberg, M. S., Waters, G., Barnes, M., & Tanenhaus, M. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, **23,** 383–404.

Shallice, T., & McCarthy, R. (1985). Phonological reading: From patterns of impairment to possible procedures. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading* (pp. 361–397). Hillsdale, NJ: Erlbaum.

Stanback, M. L. (1992). Syllable and rime patterns for teaching reading: Analysis of a frequency-based vocabulary of 17,602 words. *Annals of Dyslexia,* **42,** 196–221.

Taft, M. (1991). *Reading and the mental lexicon.* Hillsdale, NJ: Erlbaum.

Taft, M., & Radeau, M. (1995). The influence of the phonological characteristics of a language on the functional units or reading: A study in French. *Canadian Journal of Psychology,* **49,** 330–346.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language,* **26,** 608–631.

Treiman, R., (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology,* **39,** 161–181.

Treiman, R. (1992). Experimental studies of English syllabification. In W. U. Dressler, H. C. Luschutzky, O. E. Pfeiffer, & J. R. Rennison (Eds.), *Phonologica, Proceedings of the 6th International Phonology meeting.* Cambridge: Cambridge Univ. Press.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory and Cognition,* **18,** 559–567.

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: General,* **124,** 107–136.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition,* **10,** 434–442.

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

Verstaen, A., Humphreys, G. W., Olson, A., & d'Ydewalle, G. (1995). Are phonemic effects in backward masking evidence for automatic prelexical phonemic activation in visual word recognition? *Journal of Memory and Language,* **34,** 335–356.

(Received December 30, 1995)
(Revision received January 28, 1997)