Journal of Experimental Psychology:
Human Perception and Performance
2002, Vol. 28, No. 3, 661–681

0096-1523/02/$5.00 DOI: 10.1037//0096-1523.28.3.661

Orthographic Neighborhood Effects in Lexical Decision: The Effects of Nonword Orthographic Neighborhood Size

Paul D. Siakaluk and Christopher R. Sears
University of Calgary

Stephen J. Lupker
University of Western Ontario

The effects of large neighborhoods (neighborhood size) and of higher frequency neighbors (neighborhood frequency) were examined as a function of nonword neighborhood size in lexical decision tasks. According to the multiple read-out model (J. Grainger & A. M. Jacobs, 1996), neighborhood size and neighborhood frequency effects should vary systematically as a function of nonword neighborhood size. In these experiments, the nonword context was more extensively manipulated than in previous studies, providing a more complete test of the model's predictions. In addition, simulations were conducted examining the model's ability to account for the facilitatory neighborhood size and neighborhood frequency effects observed in these experiments. The results suggest that the model overestimates the role of inhibition in the orthographic processing of English words.

In the past decade, there has been a considerable amount of research on the question of whether the speed and accuracy with which a word is identified is affected by the existence of other, orthographically similar words (for a review, see Andrews, 1997). Most of this work has investigated the impact of the nature of a word's *orthographic neighborhood.* A word's orthographic neighborhood is defined as the set of different words that can be created by changing one letter of a word while maintaining letter positions (Coltheart, Davelaar, Jonasson, & Besner, 1977). For example, the words *CAKE, LAKE, BIKE,* and *BARE* are all orthographic neighbors of the word *BAKE.*

Using this definition, researchers have focused on two characteristics of a word's orthographic neighborhood. The first characteristic is the number of neighbors that a word possesses, usually referred to as the word's *neighborhood size* (or as *Coltheart's N*). The neighborhood sizes of words vary considerably, with some words (e.g., *MALE*) having more than twenty neighbors, and others (e.g., *GIRL,* with two neighbors or *IDOL,* with no neighbors) having very few. The second characteristic of interest is the existence of higher frequency neighbors in the word's orthographic neighborhood, which is usually referred to as the *neighborhood frequency.* For example, the word *PLOT* has no higher frequency neighbors. The Kucera and Francis (1967) normative frequency for *PLOT* is 37 occurrences per million words, whereas the normative frequency of its highest frequency neighbor (*PLOW*) is 12 per million. In contrast, the word *LIME* has many higher frequency neighbors. Specifically, *LIME* has a normative frequency of 13 per million, and the normative frequencies of its highest frequency neighbors, *TIME, LIKE, LIFE, LINE,* and *LIVE,* are 1,599, 1,290, 715, 298, and 177 per million, respectively.

For many models of visual word recognition, the number of neighbors and the existence of higher frequency neighbors have important processing implications (Forster, 1976; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Paap, Newsome, McDonald, & Schvaneveldt, 1982). These models assume that when a word is presented, the lexical representations of the word and its neighbors are activated, and once activated, the lexical representations of the orthographic neighbors of the word then play a role in the lexical selection (i.e., word identification) process. The precise role that orthographic neighbors play in word identification differs from model to model. As a result, considerable research has been devoted to evaluating the models' specific predictions, primarily using a lexical decision task. Unfortunately, several inconsistent findings have emerged from this body of empirical research.[1]

Paul D. Siakaluk and Christopher R. Sears, Department of Psychology, University of Calgary, Calgary, Alberta, Canada; Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, Canada.

This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant to Christopher R. Sears and a NSERC postgraduate scholarship to Paul D. Siakaluk. The experiments were conducted as part of a master's thesis by Paul D. Siakaluk under the supervision of Christopher R. Sears. Portions of these data were presented at the annual meeting of Canadian Society for Brain, Behavior, and Cognitive Science, Edmonton, Alberta, Canada (June 1999).

We are indebted to Walter van Heuven for providing us with the software we used to implement the multiple read-out model. We thank Kenneth Forster and two anonymous reviewers for their very valuable comments on a previous version of this article.

Correspondence concerning this article should be addressed to Christopher R. Sears, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada. E-mail: sears@ucalgary.ca

[1] Orthographic neighborhood effects have also been investigated in tasks other than lexical decision: the naming task (Andrews, 1989, 1992; Carreiras et al., 1997; Grainger, 1990; Peereman & Content, 1995; Sears et al., 1995), the semantic categorization task (Carreiras et al., 1997; Forster & Shen, 1996; Sears, Lupker, & Hino, 1999), and perceptual identification tasks (Carreiras et al., 1997; Grainger & Jacobs, 1996; Grainger & Segui, 1990; Sears, Lupker, & Hino, 1999; Snodgrass & Minzer, 1993; Van Heuven, Dijkstra, & Grainger, 1998). For a review, see Andrews (1997).

In what follows, we first briefly review the literature on the effects of neighborhood size and neighborhood frequency in the lexical decision task. We then describe Grainger and Jacobs's (1996) multiple read-out model, which attempts to account for the literature's inconsistencies by proposing that the nature of these effects (i.e., whether they are facilitatory, inhibitory, or nonexistent) is critically dependent on the orthographic neighborhood sizes of the nonwords used. This proposal is the major empirical focus of the present experiments, and thus, the multiple read-out model serves as the focus for much of our discussion.[2]

Effects of Neighborhood Size

In a recent review of the literature on orthographic neighborhood effects, Andrews (1997) noted that in the majority of lexical decision studies using English words, words with large neighborhoods were responded to more rapidly than were words with small neighborhoods (a facilitatory neighborhood size effect). It is important to note that this facilitatory neighborhood size effect was usually observed for only low-frequency words. Andrews (1989, 1992), for example, factorially manipulated word frequency and neighborhood size and reported that responses to low-frequency words with large neighborhoods were faster than responses to low-frequency words with small neighborhoods, whereas neighborhood size had little or no effect on the response latencies for high-frequency words. Facilitatory neighborhood size effects for low-frequency words in lexical decision tasks have also been reported by Chateau and Jared (2000), Forster and Shen (1996), Lewellen, Goldinger, Pisoni, and Greene (1993), Pollatsek, Perea, and Binder (1999), and Sears, Hino, and Lupker (1995), with the last group of investigators also reporting an interaction between word frequency and neighborhood size.

Not all investigators have reported facilitatory neighborhood size effects in the lexical decision task, however (e.g., Carreiras, Perea, & Grainger, 1997; Grainger & Jacobs, 1996; Johnson & Pugh, 1994). Grainger and Jacobs (1996) manipulated the neighborhood size of the nonwords used in a lexical decision task and reported that when the nonwords had large neighborhoods, there was no effect of neighborhood size for words (when the nonwords had small neighborhoods, however, there was a facilitatory effect of neighborhood size). Carreiras et al. (1997) intermixed small and large neighborhood nonwords in a lexical decision task and reported no effect of neighborhood size. Finally, Johnson and Pugh (1994) blocked their word and nonword stimuli by neighborhood size. That is, words and nonwords with large neighborhoods were presented in one block of trials, and words and nonwords with small neighborhoods were presented in another block of trials. Under these conditions, an inhibitory neighborhood size effect was observed (i.e., slower responses for large neighborhood words). When their word and nonword stimuli were not blocked by neighborhood size, however, they reported a trend toward a facilitatory neighborhood size effect.

Andrews (1997) argued that, taken together, the empirical findings regarding the effects of neighborhood size are fairly straightforward. When English words are used in "standard" lexical decision tasks, in which words of varying neighborhood sizes are intermixed with nonwords of varying neighborhood sizes, a facilitatory neighborhood size effect is typically observed. The inconsistent findings regarding the effects of neighborhood size emerge only when languages other than English are used (French: Grainger & Jacobs, 1996, Experiment 1B; Spanish: Carreiras et al., 1997, Experiment 2; Dutch: Van Heuven, Dijkstra, & Grainger, 1998, Experiment 3), or when the words and nonwords are blocked by neighborhood size (Johnson & Pugh, 1994).

Andrews (1997) proposed a language-specific explanation for why facilitatory effects of neighborhood size are commonly observed in English but not in French or Spanish. English has an inconsistent relationship between orthography and phonology, with vowels being more inconsistently pronounced than consonants. However, because consonants that follow a vowel better predict its correct pronunciation than consonants that precede it (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995), the word body (i.e., the orthographic rime) may play a special role when reading English words. Word bodies (and, hence, word-body neighbors) may play lesser roles in French or Spanish, because these languages have different orthographic–phonological structures (e.g., in French, final consonants are the more inconsistent component of the word; Content, 1991).

In a recent experiment, Ziegler and Perry (1998) examined the effects of body neighbors and neighborhood size for English words. When neighborhood size was controlled and the number of body neighbors was manipulated (for both word and nonword stimuli), responses to words with many body neighbors were faster than responses to words with few body neighbors. When the number of body neighbors was controlled and neighborhood size was manipulated, words with large neighborhoods were responded to slightly more slowly than words with small neighborhoods (an inhibitory neighborhood size effect).

Ziegler and Perry's (1998) results are consistent with Andrews's (1997) conjecture and suggest that the neighborhood size effect for English words may occur because words with large neighborhoods have more body neighbors than words with small neighborhoods and, further, that words with many body neighbors are processed more rapidly than words with few body neighbors. If true, this may explain why facilitatory effects of neighborhood size are consistently observed in English but not in French or Spanish. Of course, this explanation is currently based on a single experiment, so further investigation is required before any definitive conclusions can be reached.

One point to note about Ziegler and Perry's (1998) experiment is that the authors used a slightly different definition of neighbors than previously used. For example, words like *FEED* and *NEED* were considered to be neighbors of the target word *BLEED* in spite of the fact they differed in length and, hence, would not be neighbors according to the definition used by previous researchers (e.g., Coltheart et al., 1977). It is also worth noting that facilitatory neighborhood size effects have been observed with French words (Bozon & Carbonnel, 1996; Grainger & Jacobs, 1996, Experiment 1C; Mathey & Zagar, 1996; Pynte, 2000) and with Spanish words

---

[2] It should be noted that the multiple read-out model is intended to be a general model of visual word recognition and should, therefore, be able to explain performance in tasks such as perceptual identification and semantic categorization as well. A discussion of the model's ability to explain performance in tasks other than lexical decision is beyond the scope of the present article (see Carreiras et al., 1997; Grainger & Jacobs, 1996; Sears, Lupker, & Hino, 1999).

(Carreiras et al., 1997, Experiment 3). Thus, a completely language-specific account of neighborhood size effects would not appear to be a correct account.

Effects of Higher Frequency Neighbors

The second neighborhood characteristic of interest has been the frequency of the neighbors. Specifically, a number of investigators have examined the impact on response latencies for words with higher frequency neighbors (e.g., Carreiras et al., 1997; Forster & Shen, 1996; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989, 1992; Grainger & Segui, 1990; Huntsman & Lima, 1996; Perea & Pollatsek, 1998; Sears et al., 1995). Most of these studies seemed to show that lexical decision latencies to low-frequency words with higher frequency neighbors are slower than those to low-frequency words without higher frequency neighbors (usually referred to as an *inhibitory neighborhood frequency effect*). In Grainger et al.'s (1989) Experiment 1, for example, neighborhood frequency was manipulated by using words with no neighbors, words with some neighbors but none of higher frequency, words with exactly one higher frequency neighbor, and words with many higher frequency neighbors. Target word frequency was equated across these four conditions. Responses to words with at least one higher frequency neighbor were slower than responses to words with no higher frequency neighbors.

In her review of the orthographic neighborhood literature, Andrews (1997) noted that for languages other than English (e.g., French, Dutch, and Spanish), results such as Grainger et al.'s (1989) were typical (i.e., neighborhood frequency effects are usually inhibitory in a lexical decision task; see Mathey & Zagar, 2000, for an exception). Andrews also noted, however, that the effects of neighborhood frequency are less consistent when English stimuli are used. That is, some investigators have reported inhibitory neighborhood frequency effects (Huntsman & Lima, 1996; Perea & Pollatsek, 1998), whereas other investigators have reported either null or facilitatory neighborhood frequency effects (Forster & Shen, 1996; Sears et al., 1995). Thus, there is currently no consensus as to the effects of higher frequency neighbors on word identification. This lack of consensus was a major motivation behind Grainger and Jacobs's (1996) multiple read-out model, which attempts to account for both inhibitory and facilitatory neighborhood frequency effects (as well as facilitatory neighborhood size effects) in lexical decision.

The Multiple Read-Out Model

Grainger and Jacobs's (1996) multiple read-out model is based on the architecture of the interactive-activation model (McClelland & Rumelhart, 1981). When a word is presented to the model, activation spreads through a network of sublexical units and lexical units. These two types of units mutually excite one another via a reciprocal activation mechanism, which enables partially activated lexical units to eventually exceed an activation threshold as they build up activation over time. Intralevel inhibition also occurs between the lexical units. That is, the lexical units that are activated during the presentation of a word compete against one another during the lexical selection process via mutual inhibitory connections. According to the model, high-frequency words have higher resting activation levels than low-frequency words and hence can exert more inhibition on their low-frequency neighbors. This is the basic mechanism in the model that allows it to explain inhibitory neighborhood frequency effects.

The multiple read-out model is different from the interactive-activation model in that it also incorporates three variable decision criteria that influence the speed with which lexical decision responses are made. The first is the M criterion, which is sensitive to the activation of single lexical units. According to the model, when the M criterion is exceeded, lexical selection has occurred (i.e., a specific word has been identified). The second is the Σ criterion, which is sensitive to the degree of overall lexical activation. If a letter string generates enough lexical activity to exceed the current Σ criterion, then a *word* response can be made prior to lexical selection (i.e., prior to the M criterion being exceeded). The third criterion is the T criterion, which is a temporal deadline used for generating *nonword* responses. According to the model, when a letter string is presented and either the M criterion or the Σ criterion are reached before the T criterion, then a *word* response will be made; otherwise a *nonword* response will be made.

The setting of the M criterion is fixed in the model, whereas the setting of the Σ and T criteria can be strategically adjusted, on the basis of either the task instructions regarding speed and accuracy or the nature of the stimuli (e.g., the overlap in the neighborhood sizes of the word and nonword stimuli). The particular setting of the Σ criterion determines whether positive responses are based more on lexical selection or global lexical activity. Specifically, when the Σ criterion is set relatively high, the M criterion will usually be reached first, and most *word* responses will occur following lexical selection. Conversely, when the Σ criterion is set relatively low, the Σ criterion will usually be reached before the M criterion, and most *word* responses will be made on the basis of global lexical activity, prior to lexical selection.

With regard to orthographic neighborhood effects, the major assumptions in the model are (a) facilitatory neighborhood size effects (and any facilitatory neighborhood frequency effects) in lexical decision do not actually arise as a result of the lexical selection process (i.e., due to the M criterion being reached), but instead occur when participants use the Σ criterion for responding, and (b) the inhibitory neighborhood frequency effect is a true lexical selection effect, resulting from the intralevel competitive processes that occur prior to the M criterion being reached.

To test these assumptions, Grainger and Jacobs (1996) conducted two lexical decision experiments in which the neighborhood size of the nonwords was varied. In these experiments, for the word stimuli, neighborhood size and neighborhood frequency were manipulated by using (a) words with few neighbors and none of higher frequency, (b) words with few neighbors, with one of higher frequency, (c) words with many neighbors, of which one was of higher frequency, and (d) words with many neighbors, of which two or more were of higher frequency. (All of the words in each condition were of low frequency.) The critical comparisons were between Conditions 1 and 2 (the effect of one higher frequency neighbor), Conditions 2 and 3 (the effect of neighborhood size), and Conditions 3 and 4 (the effect of number of higher frequency neighbors).

In Grainger and Jacobs's (1996) Experiment 1C, all of the nonwords had small neighborhoods. According to the multiple read-out model, in this situation, the Σ criterion should generally be set relatively low in comparison to the M criterion because lexical activation will be a reliable cue as to whether or not a stimulus is a word (because the word stimuli will generate, on average, more lexical activity than the nonword stimuli). When the Σ criterion does drive responding, words with large neighborhoods will be easier to distinguish from the nonwords than words with small neighborhoods, because words with large neighborhoods will produce more lexical activity than words with small neighborhoods. Consequently, a facilitatory neighborhood size effect should be observed (Condition 2 vs. Condition 3). Further, according to Grainger and Jacobs, as some of the words with small neighborhoods and nonwords with small neighborhoods will not be distinguishable from one another on the basis of lexical activity, the M criterion should be occasionally used for responding. Consequently, an inhibitory neighborhood frequency effect should also be observed (Condition 1 vs. Condition 2). This was the case in Grainger and Jacobs's Experiment 1C, and simulations with the model indicated it was successful in accounting for these effects.

In Grainger and Jacobs's (1996) Experiment 1B, all of the nonwords had large neighborhoods, and in this situation, according to the model, the Σ criterion should be set relatively high, because the degree of lexical activation will not be useful for distinguishing the words from the nonwords (i.e., nonwords with large neighborhoods will generate a great deal of lexical activity, and thus, it would be difficult to distinguish them from the words on the basis of this activity). Consequently, the participant must wait until lexical selection is completed before making a response (i.e., the M criterion must be exceeded), and thus, a facilitatory neighborhood size effect should not be observed (because a facilitatory neighborhood size effect will occur only when the Σ criterion is used for responding). In addition, the inhibitory neighborhood frequency effect should be larger in this situation relative to that observed when all of the nonwords had small neighborhoods (i.e., their Experiment 1C), because more of the responses should be based on the M criterion being exceeded. This was indeed the case, and, once again, the model was successful in simulating the results.

Grainger and colleagues (Carreiras et al., 1997; Grainger & Jacobs, 1996) have proposed that the multiple read-out model can account for many of the conflicting findings in the literature regarding the effects of neighborhood size and neighborhood frequency. Specifically, reports of facilitatory neighborhood size effects (Andrews, 1989, 1992; Forster & Shen, 1996; Sears et al., 1995) are explained by assuming that the participants in these experiments made extensive use of the Σ criterion when responding, whereas reports of null and inhibitory neighborhood size effects (e.g., Carreiras et al., 1997; Johnson & Pugh, 1994) are explained by assuming that these participants relied more on the M criterion than on the Σ criterion when responding. Similarly, reports of null or facilitatory neighborhood frequency effects (Forster & Shen, 1996; Sears et al., 1995) can be explained by assuming that (a) words with higher frequency neighbors produce more lexical activation than words without higher frequency neighbors, and (b) the participants in these experiments relied almost entirely on the Σ criterion when responding.

Experiment 1

The purpose of the present research was to provide a thorough examination of the effects of nonword neighborhood size on orthographic neighborhood effects in the lexical decision task and to evaluate the multiple read-out model's ability to account for such effects. As in Grainger and Jacobs's (1996) experiments, in the present Experiments 1A–1D, the neighborhood sizes of the nonwords used in lexical decision tasks were manipulated to test the multiple read-out model's account of neighborhood size and neighborhood frequency effects in these tasks. The present experiments differed from those of Grainger and Jacobs, however, in several important ways.

First, in the present experiments, word frequency (high or low), neighborhood size (small or large), and neighborhood frequency (no higher frequency neighbors or higher frequency neighbors) were factorially manipulated. (The same set of words was used in each of the experiments.) This design was expected to better clarify how neighborhood size and neighborhood frequency affect the processing of both high-frequency and low-frequency words and to allow for tests of any interactions between these factors.

The word frequency manipulation also allowed us to gauge the extent to which lexical selection was involved during responding in each experiment. According to the multiple read-out model, word frequency effects arise because high-frequency words have higher resting activation levels than low-frequency words and thus will exceed the M criterion quite rapidly during processing. As such, differences in the magnitude of the word frequency effect between the experiments can be used to gauge the extent to which responses were based on the M criterion (i.e., lexical selection being achieved), with larger word frequency effects reflecting increased use of the M criterion. It is important to note that because the inhibitory neighborhood frequency effect is also assumed to be a lexical selection effect (resulting from the intralevel competitive processes that occur prior to the M criterion being exceeded), we expected a direct relation between the magnitude of the word frequency effect and the magnitude of the neighborhood frequency effect. Specifically, the word frequency effect and the neighborhood frequency effect were expected to be positively correlated (i.e., a larger word frequency effect would correspond with a larger inhibitory neighborhood frequency effect, and vice versa). Conversely, the word frequency effect and the neighborhood size effect were expected to be negatively correlated (i.e., smaller word frequency effects would correspond with larger neighborhood size effects, and vice versa).

Second, in Experiments 1A–1D, the nonword context was more extensively manipulated than in Grainger and Jacobs's (1996) experiments. Specifically, Grainger and Jacobs used two nonword contexts (the nonwords had either small neighborhoods or large neighborhoods), whereas in the present experiments, there were four contexts: The nonwords had no neighbors, they had small neighborhoods, they had large neighborhoods, or they were matched to the words on neighborhood size. This more extensive manipulation of the nonword context allowed for a more comprehensive test of the model's predictions regarding the effects of orthographic neighbors in the lexical decision task.

In Experiment 1A, the nonwords had no neighbors (e.g., *NALB*). According to the multiple read-out model, in this experiment, the words will generate much more lexical activity than the nonwords, and, therefore, participants will be able to use the degree of lexical activation as a basis for responding (i.e., if the stimulus produces a great deal of lexical activity, then it is probably a word). Participants should have thus set their Σ criterion quite low, which means that it and not the M criterion would have been used for most (if not all) of the *word* responses. In this situation, the model predicts that both the neighborhood size effect and the neighborhood frequency effect will be facilitatory, because words with large neighborhoods and words with higher frequency neighbors will generate more lexical activity than words with small neighborhoods and words with no higher frequency neighbors. The model further predicts that the facilitatory effects of neighborhood size should be greatest in this experiment, because more responses should be based on the Σ criterion in this experiment than in any of the other experiments. In addition, because few of the responses would be based on the M criterion (i.e., lexical selection being achieved), the word frequency effect was expected to be fairly small, because the model assumes that only the lexical selection process itself is sensitive to word frequency.

In Experiment 1B, the nonwords had small neighborhoods (e.g., *GRUN,* with two neighbors). According to the model, in this experiment, because the large neighborhood words will generate more lexical activity than the small neighborhood words and nonwords, participants will often be using their Σ criterion rather than their M criterion. Thus, a facilitatory neighborhood size effect should be observed. Because the words with small neighborhoods and the nonwords with small neighborhoods will be difficult to distinguish from one another via the Σ criterion, however, the model further assumes that many responses to those words will be based on the M criterion, in which case an inhibitory neighborhood frequency effect should be observed. Consequently, at least for the low-frequency words, there should be an interaction between neighborhood size and neighborhood frequency—the words with small neighborhoods should exhibit an inhibitory neighborhood frequency effect, whereas the words with large neighborhoods should not. In addition, because of the greater involvement of the M criterion in responding, the word frequency effect should be larger than that observed in Experiment 1A. Note that this experiment was similar to Grainger and Jacobs's (1996) Experiment 1C and thus serves as an attempted replication of their findings with a different and more extensive set of stimuli (in particular, a stimulus set in which neighborhood size and neighborhood frequency were manipulated, allowing for a test of their predicted interaction).

In Experiment 1C, the words and the nonwords were matched on neighborhood size. In this situation, according to the model, because the degree of lexical activity generated by the words and the nonwords will be very similar, it should not be possible to reliably distinguish any of the words from any of the nonwords on the basis of global lexical activation. The model therefore assumes that most (if not all) of the *word* responses should be made using the M criterion (i.e., responses should be made following unique word identification) and thus predicts an inhibitory neighborhood frequency effect (due to the use of the M criterion) and a null neighborhood size effect (as the Σ criterion should not be used). The model further predicts that the inhibitory effects of neighborhood frequency should be larger in this experiment than in Experiment 1B, because more responses should be based on the M criterion. Similarly, the model predicts that a larger word frequency effect should be observed in this experiment, again because more responses should be based on the M criterion in this experiment than in Experiments 1A and 1B.

In Experiment 1D, the nonwords had large neighborhoods (e.g., *DAST*, with 13 neighbors), as was the case in Grainger and Jacobs's (1996) Experiment 1B. Recall that, according to the model, nonwords with large neighborhoods will generate a great deal of lexical activity, and thus, it would be difficult to distinguish them from the words on the basis of this activity. As a consequence, participants will not be able to use the degree of lexical activity as a basis for responding. Participants should thus generally set their Σ criterion somewhat high, and because virtually all of the responses will be made using the M criterion, the model predicts that an inhibitory neighborhood frequency effect should be observed. Further, the inhibitory neighborhood frequency effect in this experiment should be larger than that observed in Experiment 1B, because of the increased use of the M criterion in this experiment. Similarly, because of the increased use of the M criterion relative to Experiments 1A and 1B, the word frequency effect should also be larger in this experiment. The model also predicts that the neighborhood size effect should be minor or nonexistent, because the Σ criterion should be used only occasionally for responding. Note that the model's predictions for this experiment are essentially the same as those of Experiment 1C, because in both cases, participants will not be able to use the degree of lexical activity as a basis for responding. Thus, in addition to this experiment serving as an attempted replication of Grainger and Jacobs's Experiment 1B, a comparison of the results of Experiments 1D and 1C further tests the multiple read-out model's assumptions regarding the relative use of the Σ criterion and the M criterion in different nonword contexts.

Finally, note that in all of the experiments, efforts were made to keep the error rates fairly low. This was accomplished by instructing participants to give preference to accuracy over speed, by instructing participants to keep their error rates below 5%, and by providing error feedback (percentage of errors) after every block of 60 trials. The concern here was motivated in part by Carreiras et al.'s (1997) observation that some of the previous experiments that have produced null or facilitatory effects of neighborhood frequency (e.g., Forster & Shen, 1996; Sears et al., 1995) had higher error rates than some of the experiments that have produced inhibitory effects of neighborhood frequency (e.g., Carreiras et al., 1997; Huntsman & Lima, 1996): The idea is that high error rates suggest that participants were relying more on the Σ criterion than on the M criterion for responding (even when this is not an optimal strategy given the nonword context), in which case the likelihood of observing an inhibitory neighborhood frequency effect would be greatly reduced. Our efforts to keep error rates low were successful, and thus, error rates should not be a concern in the interpretation of our latency results.

General Method

Participants. There were 160 undergraduate students from the University of Calgary who participated in the experiments: 40 participants in each of the four experiments. All were native English speakers and reported that they had normal or corrected-to-normal vision. None of these individuals participated in more than one experiment.

Word stimuli. The complete set of experimental words used in Experiments 1A, 1B, 1C, and 1D is presented in the Appendix, and the descriptive statistics for these stimuli are listed in Table 1.

Half of the words in each of the eight conditions of the experiment were four letters in length, and the other half of the words were five letters in length. Three factors were manipulated. The first factor was word frequency. Half of the words were high-frequency words, with a mean Kucera

Table 1

Mean Word Frequency (Occurrences per Million Words) and Neighborhood Size for the Word Stimuli Used in Experiments 1A–1D

	Low-frequency words		High-frequency words	
Stimulus characteristic	Small N	Large N	Small N	Large N
No higher frequency neighbors				
Frequency	19.7	24.4	113.0	105.8
N	3.3	8.6	3.1	9.5
NF	8.2	13.9	28.1	58.0
Higher frequency neighbors				
Frequency	20.2	20.4	98.3	105.3
N	3.5	10.2	3.3	10.8
NF	288.1	315.5	261.7	328.9

Note. N = neighborhood size; NF = average frequency of the highest frequency neighbor.

and Francis (1967) normative frequency (occurrences per million words) of 105.6 (range = 52–231), and the other half were low-frequency words, with a mean normative frequency of 21.1 (range = 1–48).

The second factor manipulated was neighborhood size. Half of the words had small neighborhoods (i.e., at least one neighbor but no more than five neighbors); these words had a mean neighborhood size of 3.3. The other half of the words had large neighborhoods (i.e., at least six neighbors; range = 6–18); these had a mean neighborhood size of 9.7. To be considered a neighbor of a target word, a word had to appear either in the Kucera and Francis (1967) norms or in an 80,000-word, computer-based dictionary.

The third factor manipulated was neighborhood frequency: the presence or absence of higher frequency neighbors in a word's orthographic neighborhood. Half of the words had at least one neighbor of higher frequency, and the other half had no neighbors that were higher in frequency. For the high-frequency words with higher frequency neighbors, the mean Kucera and Francis (1967) normative frequency of the highest frequency neighbor of each word was 295.3. For the low-frequency words with higher frequency neighbors, the mean Kucera and Francis normative frequency of the highest frequency neighbor of each word was 301.8. Finally, for both the high-frequency and the low-frequency words with no higher frequency neighbors, the mean frequency of the highest frequency neighbor of each word was substantially lower than the mean target frequency.

Nonword stimuli. Four different sets of nonword stimuli were created. All of the nonwords were orthographically legal and pronounceable. In Experiment 1A, the nonwords had no orthographic neighbors. In Experiment 1B, the nonwords had small neighborhoods (range = 1–5 neighbors; mean neighborhood size = 3.5). The mean neighborhood size of these nonwords was essentially the same as the mean neighborhood size of the small neighborhood words (3.3). In Experiment 1C, the nonwords were matched closely to the words in neighborhood size. More specifically, as noted, for the small neighborhood words, the mean neighborhood size was 3.3, and for the small neighborhood nonwords, the mean neighborhood size was 3.4 (range = 1–5 neighbors). The mean neighborhood sizes of the large neighborhood words and the large neighborhood nonwords were 9.8 and 9.6, respectively (for both, range = 6–18 neighbors). (The overall mean neighborhood size of the nonwords used in Experiment 1C was 6.5.) In Experiment 1D, the nonwords had large neighborhoods (range = 6–21 neighbors; mean neighborhood size = 9.6), which was virtually identical to the mean neighborhood size of the large neighborhood words.

Apparatus and procedure. Stimuli were presented on a color video graphics array monitor driven by a Pentium-class microcomputer. The presentation of stimuli was synchronized with the vertical retrace rate of the monitor (14 ms), and response latencies were measured to the nearest millisecond. At a viewing distance of 50 cm, the stimuli subtended a visual angle of approximately 1.1 degrees.

Each trial was initiated by a 1-s, 2000-Hz warning tone, after which a fixation point appeared at the center of the video monitor. The fixation point was presented for 1 s, and then was replaced by a word or nonword stimulus (presented in uppercase letters). Participants indicated the lexicality of stimuli (*word* or *nonword*) by pressing one of two buttons on a response box. The participant's response terminated the stimulus display, and the next trial was initiated after a timed interval of 2 s.

Each participant completed 16 practice trials prior to the collection of data. The practice stimuli consisted of eight words (four of low frequency and four of high frequency) and eight orthographically legal and pronounceable nonwords. The nonwords used in the practice trials were representative of the nonwords presented in the experimental trials (e.g., the eight nonwords for the practice trials of Experiment 1A had no neighbors). (These practice stimuli were not used in the actual experiment, and the data from these practice trials were not analyzed.) Following the practice trials, the participants were provided with feedback as to the mean latency and accuracy of their responses (percentage of errors), and during the experimental trials, this information was presented every 60 trials. Participants were instructed to respond as quickly as possible while keeping their error rate below 5%.

Design. A 2 (word frequency: high or low) × 2 (neighborhood size: small or large) × 2 (neighborhood frequency: no higher frequency neighbors or higher frequency neighbors) factorial design was used for each of the experiments. There were 26 words in each of the eight stimulus conditions, for a total of 208 words. There were also 208 nonwords presented in each experiment (half of them four letters in length and the other half five letters in length), for a total of 416 trials. The order in which the stimuli were presented in the experiments was randomized separately for each participant.

For the word data, response latencies and error rates from each experiment were submitted to a 2 (word frequency: high or low) × 2 (neighborhood size: small or large) × 2 (neighborhood frequency: no higher frequency neighbors or higher frequency neighbors) repeated measures factorial analysis of variance (ANOVA). Both subject (F_s) and item (F_i) analyses were performed.[3]

---

[3] Many of the item analyses reported in the present article were not statistically significant. We do not regard this as a particularly important issue. Although Clark (1973) has argued that items, as well as subjects, should be considered as a random factor in these types of analyses, it is seldom the case that the selection of items is ever random in any sense of the term. That is, typically, the items used in these types of experiments have been selected because they satisfied an extensive set of criteria, which is certainly the case in the experiments reported here (e.g., see Table 1). Consequently, as Wike and Church (1976) and others (Cohen, 1976; Keppel, 1976; see also Raaijmakers, Schrijnemakers, & Gremmen, 1999) have argued, item analyses would clearly be inappropriate in the present situation for a number of reasons, not the least of which is their strong negative bias (i.e., when items have not been selected randomly, the statistical power of item analyses is reduced because of a greatly deflated alpha value). Any concerns about the generalizability of the results across items are addressed by the significant neighborhood size and the neighborhood frequency effects in item analyses when the stimuli from Experiments 1A and 1B, Experiments 1A and 1C, and Experiments 1B and 1C were combined. It should also be noted that Grainger and Jacobs (1996) did not report any item analyses for their lexical decision experiments, so it is not known which of their results would have been significant in item analyses.

Results

The mean response latencies of correct responses and the mean error rates in Experiments 1A, 1B, 1C, and 1D are shown in Tables 2, 3, 4, and 5, respectively. Table 6 lists the mean response latencies and error rates for the words and the nonwords in each of the four experiments. Table 7 lists the mean word frequency, neighborhood size, and neighborhood frequency effect in each experiment.

Nonwords with no neighbors (Experiment 1A). In the analysis of the response latencies, there was a significant main effect of word frequency, $F_s(1, 39) = 68.53$, $p < .001$, $MSE = 786.13$, and $F_i(1, 200) = 36.16$, $p < .001$, $MSE = 1{,}004.59$, as responses to high-frequency words were an average of 26 ms faster than responses to low-frequency words. The main effect of neighborhood size was significant, $F_s(1, 39) = 20.95$, $p < .001$, $MSE = 234.74$, and $F_i(1, 200) = 3.71$, $p = .05$, $MSE = 1{,}004.59$, as was the main effect of neighborhood frequency, $F_s(1, 39) = 16.14$, $p < .001$, $MSE = 361.91$, and $F_i(1, 200) = 4.38$, $p < .05$, $MSE = 1{,}004.59$. Responses to words with large neighborhoods were an average of 8 ms faster than responses to words with small neighborhoods, and responses to words with higher frequency neighbors were an average of 8.5 ms faster than responses to words with no higher frequency neighbors.

The interaction between word frequency and neighborhood size was not significant, $F_s(1, 39) = 2.63$, $p > .10$, $MSE = 205.53$, and $F_i < 1$, as both the low-frequency words and the high-frequency words exhibited a facilitatory neighborhood size effect, $F_s(1, 39) = 27.86$, $p < .001$, $MSE = 156.54$, and $F_i(1, 100) = 2.28$, $p > .10$, $MSE = 1{,}456.23$, and $F_s(1, 39) = 3.87$, $p = .05$, $MSE = 283.72$, and $F_i(1, 100) = 1.49$, $p > .10$, $MSE = 552.96$, respectively. The interaction between word frequency and neighborhood frequency was significant, $F_s(1, 39) = 4.68$, $p < .05$, $MSE = 262.38$, and $F_i < 1$. For low-frequency words, the neighborhood frequency effect was facilitatory—responses to words with higher frequency neighbors were an average of 12.5 ms faster than responses to words with no higher frequency neighbors, $F_s(1, 39) = 18.66$, $p < .001$, $MSE = 332.97$, and $F_i(1, 100) = 3.15$, $p = .07$, $MSE = 1{,}456.23$. For high-frequency words, responses to words with higher frequency neighbors were an average of 4.5 ms faster than responses to words with no higher frequency neighbors, $F_s(1, 39) = 2.94$, $p = .09$, $MSE = 291.32$, and $F_i(1, 100) = 1.23$, $p > .10$, $MSE = 552.96$. Thus, the interaction occurred because the facilitatory neighborhood frequency effect was larger for the low-frequency words than for the high-frequency words. No other interactions were significant (all Fs < 1).

Table 3
Mean Response Latencies (in Milliseconds), Standard Errors, and Error Rates (Percentages) in Experiment 1B

	Neighborhood size	
Neighborhood frequency	Small	Large
Low-frequency words		
No HF N	575 (7.7, 5.7)	562 (7.5, 4.6)
HF N	564 (6.9, 5.2)	548 (6.3, 3.8)
High-frequency words		
No HF N	525 (6.2, 1.9)	519 (6.2, 1.4)
HF N	519 (6.1, 1.1)	521 (5.7, 1.1)

Note. HF N = higher frequency neighbors. Standard errors and error rates, respectively, appear in parentheses.

In the analysis of error rates, the main effect of word frequency was significant, $F_s(1, 39) = 27.36$, $p < .001$, $MSE = 5.49$, and $F_i(1, 200) = 12.67$, $p < .001$, $MSE = 7.70$, as was the main effect of neighborhood size, $F_s(1, 39) = 13.49$, $p < .01$, $MSE = 8.91$, and $F_i(1, 200) = 10.15$, $p < .01$, $MSE = 7.70$. Participants made fewer errors to high-frequency words than to low-frequency words (1.2% vs. 2.6%), and fewer errors to words with large neighborhoods than to words with small neighborhoods (1.3% vs. 2.5%). The main effect of neighborhood frequency was also significant, $F_s(1, 39) = 4.31$, $p < .05$, $MSE = 6.70$, and $F_i(1, 200) = 2.44$, $p > .10$, $MSE = 7.70$. Participants made fewer errors to words with higher frequency neighbors than to words with no higher frequency neighbors (1.6% vs. 2.2%).

As was the case in the response latency analysis, the Word Frequency × Neighborhood Frequency interaction was significant, $F_s(1, 39) = 10.21$, $p < .01$, $MSE = 6.20$, and $F_i(1, 200) = 5.34$, $p < .05$, $MSE = 7.70$. For the low-frequency words, words with higher frequency neighbors were responded to more accurately

Table 2
Mean Response Latencies (in Milliseconds), Standard Errors, and Error Rates (Percentages) in Experiment 1A

	Neighborhood size	
Neighborhood frequency	Small	Large
Low-frequency words		
No HF N	517 (8.1, 4.2)	509 (7.7, 2.5)
HF N	507 (8.3, 2.7)	494 (7.8, 1.0)
High-frequency words		
No HF N	486 (6.5, 1.6)	480 (7.2, 0.6)
HF N	481 (6.9, 1.7)	476 (7.1, 1.1)

Note. HF N = higher frequency neighbors. Standard errors and error rates, respectively, appear in parentheses.

Table 4
Mean Response Latencies (in Milliseconds), Standard Errors, and Error Rates (Percentages) in Experiment 1C

	Neighborhood size	
Neighborhood frequency	Small	Large
Low-frequency words		
No HF N	613 (11.2, 5.6)	593 (11.7, 4.1)
HF N	598 (12.6, 3.4)	579 (9.4, 2.8)
High-frequency words		
No HF N	544 (9.4, 1.9)	542 (9.8, 2.1)
HF N	543 (9.1, 0.9)	541 (10.0, 1.1)

Note. HF N = higher frequency neighbors. Standard errors and error rates, respectively, appear in parentheses.

Table 5
Mean Response Latencies (in Milliseconds), Standard Errors, and Error Rates (Percentages) in Experiment 1D

Neighborhood frequency	Neighborhood size: Small	Neighborhood size: Large
Low-frequency words		
No HF N	615 (10.4, 4.5)	609 (9.2, 4.0)
HF N	600 (12.0, 3.9)	602 (11.5, 4.1)
High-frequency words		
No HF N	557 (9.2, 2.2)	568 (10.9, 2.8)
HF N	552 (9.2, 0.9)	564 (10.0, 2.0)

Note. HF N = higher frequency neighbors. Standard errors and error rates, respectively, appear in parentheses.

than words with no higher frequency neighbors (1.8% vs. 3.3%), $F_s(1, 39) = 14.88$, $p < .001$, $MSE = 5.97$, and $F_i(1, 100) = 5.33$, $p < .05$, $MSE = 10.84$. For the high-frequency words, there was no effect of neighborhood frequency (both *F*s < 1). No other interactions were significant (all *F*s < 1).

Nonwords with small neighborhoods (Experiment 1B). The main effect of word frequency was significant, $F_s(1, 39) = 161.22$, $p < .001$, $MSE = 832.89$, and $F_i(1, 200) = 54.70$, $p < .001$, $MSE = 1{,}792.09$. Responses to high-frequency words were an average of 41 ms faster than responses to low-frequency words. The main effect of neighborhood size was significant, $F_s(1, 39) = 11.09$, $p < .01$, $MSE = 449.08$, and $F_i(1, 200) = 2.22$, $p > .10$, $MSE = 1{,}792.09$, as was the main effect of neighborhood frequency, $F_s(1, 39) = 5.58$, $p < .05$, $MSE = 736.35$, and $F_i(1, 200) = 1.77$, $p > .10$, $MSE = 1{,}792.09$. Responses to words with large neighborhoods were an average of 8 ms faster than responses to words with small neighborhoods, and responses to words with higher frequency neighbors were an average of 7 ms faster than responses to words with no higher frequency neighbors.

The interaction between word frequency and neighborhood size was significant, $F_s(1, 39) = 6.59$, $p < .05$, $MSE = 476.66$, and $F_i(1, 200) = 1.42$, $p > .10$, $MSE = 1{,}792.09$. For the low-frequency words, responses to words with large neighborhoods were an average of 14.5 ms faster than responses to words with small neighborhoods, $F_s(1, 39) = 13.36$, $p < .01$, $MSE = 600.17$, and $F_i(1, 100) = 2.31$, $p > .10$, $MSE = 2{,}786.34$. For the high-frequency words, there was no neighborhood size effect (both *F*s < 1).

Table 6
Summary of Mean Response Latencies (in Milliseconds) and Error Rates (Percentages) in Experiments 1A–1D

Experiment	All words	LF words	HF words	Nonwords
1A	493 (1.9)	506 (2.6)	480 (1.2)	542 (2.7)
1B	541 (3.1)	562 (4.8)	521 (1.3)	616 (4.7)
1C	569 (2.7)	595 (3.9)	542 (1.5)	658 (4.3)
1D	583 (3.0)	606 (4.1)	560 (1.9)	681 (4.9)

Note. LF = low frequency; HF = high frequency. Error rates appear in parentheses.

Table 7
Summary of Word Frequency, Neighborhood Size (N), and Neighborhood Frequency (NF) Effects in Experiments 1A–1D

Experiment	Type of effect: Word frequency	Type of effect: N	Type of effect: NF
Low- and high-frequency words			
1A	26	8	8
1B	41	7	8
1C	53	10	7
1D	46	+4	7
Low-frequency words			
1A		10	12
1B		14	12
1C		19	14
1D		2	11
High-frequency words			
1A		5	4
1B		2	2
1C		2	1
1D		+11	4

Note. The neighborhood size effect was calculated as the difference (in milliseconds) between the words with small neighborhoods and the words with large neighborhoods. The neighborhood frequency effect was calculated as the difference between the words without higher frequency neighbors and the words with higher frequency neighbors. Note that decimal values have been truncated. All effects were facilitatory except where indicated by a plus sign; these effects were inhibitory.

The interaction between word frequency and neighborhood frequency was also significant, $F_s(1, 39) = 6.05$, $p < .05$, $MSE = 376.69$, and $F_i(1, 200) = 1.06$, $p > .10$, $MSE = 1{,}792.09$. For the low-frequency words, responses to words with higher frequency neighbors were an average of 12.5 ms faster than responses to words without higher frequency neighbors, $F_s(1, 39) = 8.49$, $p < .01$, $MSE = 736.91$, and $F_i(1, 100) = 1.79$, $p > .10$, $MSE = 2{,}786.64$. For the high-frequency words, there was no effect of neighborhood frequency (both *F*s < 1). The interaction between neighborhood size and neighborhood frequency was not significant (both *F*s < 1), nor was the three-way interaction, $F_s(1, 39) = 1.31$, $p > .10$, $MSE = 387.67$, and $F_i < 1$.

The main effect of word frequency was significant in the error analysis, $F_s(1, 39) = 53.26$, $p < .001$, $MSE = 18.00$, and $F_i(1, 200) = 30.03$, $p < .001$, $MSE = 20.75$. Participants made fewer errors to high-frequency words than to low-frequency words (1.3% vs. 4.8%). The main effect of neighborhood size was marginally significant, $F_s(1, 39) = 3.61$, $p = .06$, $MSE = 13.10$, and $F_i(1, 200) = 1.48$, $p > .10$, $MSE = 20.75$. Participants generally made fewer errors to words with large neighborhoods than to words with small neighborhoods (2.7% vs. 3.4%). The main effect of neighborhood frequency was not significant, $F_s(1, 39) = 2.51$, $p > .10$, $MSE = 10.60$, and $F_i < 1$, nor were any of the interactions (all *p*s > .15).

Nonwords with small and large neighborhoods (Experiment 1C). The mean response latency and percentage of errors for the small neighborhood nonwords were 641 ms and 3.3%, respectively. The mean response latency and percentage of errors for the large neighborhood nonwords were 675 ms and 5.4%, respectively. Both the latency difference, $F_s(1, 39) = 70.64$, $p < .001$, $MSE = 350.61$, and $F_i(1, 206) = 13.62$, $p < .001$, $MSE = 4,549.14$, and the accuracy difference, $F_s(1, 39) = 31.05$, $p < .001$, $MSE = 2.52$, and $F_i(1, 206) = 6.21$, $p < .05$, $MSE = 34.16$, were significant.

In the analysis of the word data, there was a significant main effect of word frequency, $F_s(1, 39) = 224.97$, $p < .001$, $MSE = 1,025.07$, and $F_i(1, 200) = 88.03$, $p < .001$, $MSE = 1,793.49$. Responses to high-frequency words were an average of 53 ms faster than responses to low-frequency words. The main effect of neighborhood size was significant, $F_s(1, 39) = 7.73$, $p < .01$, $MSE = 1,225.02$, and $F_i(1, 200) = 3.07$, $p = .08$, $MSE = 1,793.49$, as was the main effect of neighborhood frequency, $F_s(1, 39) = 6.07$, $p < .05$, $MSE = 807.77$, and $F_i(1, 200) = 2.49$, $p > .10$, $MSE = 1,793.49$, respectively. Responses to words with large neighborhoods were an average of 10 ms faster than responses to words with small neighborhoods, and responses to words with higher frequency neighbors were an average of 7 ms faster than responses to words with no higher frequency neighbors.

There was a significant interaction between word frequency and neighborhood size, $F_s(1, 39) = 10.38$, $p < .01$, $MSE = 617.59$, and $F_i(1, 200) = 2.39$, $p > .10$, $MSE = 1,793.49$. For the low-frequency words, responses to words with large neighborhoods were an average of 19.5 ms faster than responses to words with small neighborhoods, $F_s(1, 39) = 13.29$, $p < .01$, $MSE = 1,183.01$, and $F_i(1, 100) = 3.59$, $p = .06$, $MSE = 2,719.49$. For the high-frequency words, the neighborhood size effect was only 2 ms (both Fs < 1).

The interaction between word frequency and neighborhood frequency was not significant, $F_s(1, 39) = 2.72$, $p = .10$, $MSE = 1,317.84$, and $F_i(1, 200) = 1.66$, $p > .10$, $MSE = 1,793.49$. However, an examination of Table 4 suggests that only the low-frequency words exhibited a neighborhood frequency effect, which was confirmed in separate analyses of the low-frequency and high-frequency words. Specifically, for the low-frequency words, responses to words with higher frequency neighbors were an average of 14.5 ms faster than responses to words with no higher frequency neighbors, $F_s(1, 39) = 5.22$, $p < .05$, $MSE = 1,616.84$, and $F_i(1, 100) = 2.71$, $p = .10$, $MSE = 2,719.49$. For the high-frequency words, the neighborhood frequency effect was only 1 ms (both Fs < 1). No other interactions were significant (all Fs < 1).[4]

In the analysis of error rates, the main effect of word frequency was significant, $F_s(1, 39) = 44.76$, $p < .001$, $MSE = 11.17$, and $F_i(1, 200) = 17.47$, $p < .001$, $MSE = 18.60$, as participants made fewer errors to high-frequency words than to low-frequency words (1.5% vs. 3.9%). The main effect of neighborhood frequency was also significant, $F_s(1, 39) = 12.35$, $p < .01$, $MSE = 11.74$, and $F_i(1, 100) = 5.07$, $p < .05$, $MSE = 18.60$. Fewer errors were made to words with higher frequency neighbors than to words with no higher frequency neighbors (2.0% vs. 3.4%). The main effect of neighborhood size was not significant, $F_s(1, 39) = 1.90$, $p > .10$, $MSE = 7.87$, and $F_i < 1$.

The interaction between word frequency and neighborhood size was marginally significant, $F_s(1, 39) = 3.28$, $p = .07$, $MSE = 9.53$, and $F_i\ (1, 200) = 1.09$, $p > .10$, $MSE = 18.60$. For the low-frequency words, fewer errors were made to words with large neighborhoods than to words with small neighborhoods (3.4% vs. 4.5%), $F_s(1, 39) = 3.63$, $p = .06$, $MSE = 12.32$, and $F_i < 1$. For the high-frequency words, neighborhood size had no effect on error rates (both Fs < 1). The interaction between word frequency and neighborhood frequency was not significant, $F_s(1, 39) = 1.79$, $p > .10$, $MSE = 6.62$, and $F_i < 1$, nor were any of the other interactions (all Fs < 1).

Nonwords with large neighborhoods (Experiment 1D). There was a significant main effect of word frequency in the response latency analysis, $F_s(1, 39) = 176.02$, $p < .001$, $MSE = 969.26$, and $F_i(1, 200) = 60.01$, $p < .001$, $MSE = 2,058.28$, as responses to high-frequency words were an average of 46 ms faster than responses to low-frequency words. The main effect of neighborhood size was marginally significant, $F_s(1, 39) = 3.25$, $p = .07$, $MSE = 558.08$, and $F_i < 1$. Responses to words with large neighborhoods were an average of 4 ms slower than responses to words with small neighborhoods. The main effect of neighborhood frequency was significant, $F_s(1, 39) = 7.95$, $p < .01$, $MSE = 640.39$, and $F_i(1, 200) = 1.80$, $p > .10$, $MSE = 2,058.28$. Responses to words with higher frequency neighbors were an average of 7 ms faster than responses to words without higher frequency neighbors. The interaction between word frequency and neighborhood frequency was not significant (both Fs < 1).

The interaction between word frequency and neighborhood size was marginally significant, $F_s(1, 39) = 4.06$, $p = .05$, $MSE = 896.58$, and $F_i(1, 200) = 1.24$, $p > .10$, $MSE = 2,058.28$. An inspection of Table 5 reveals that for the high-frequency words, there appeared to be an inhibitory neighborhood size effect, whereas for the low-frequency words, neighborhood size appeared to have no effect on response latencies. These observations were confirmed in separate analyses of the high-frequency and low-frequency words. For the high-frequency words, words with large neighborhoods were responded to more slowly than words with small neighborhoods, $F_s(1, 39) = 9.59$, $p < .01$, $MSE = 552.33$, and $F_i(1, 100) = 3.62$, $p = .06$, $MSE = 975.34$. For the low-frequency words, there was no effect of neighborhood size (both Fs < 1). No other interactions were significant (all Fs < 1).

In the analysis of error rates, the main effect of word frequency was significant, $F_s(1, 39) = 32.87$, $p < .001$, $MSE = 11.14$, and $F_i(1, 200) = 11.96$, $p < .01$, $MSE = 19.90$. Participants made fewer errors to high-frequency words than to low-frequency words (1.9% vs. 4.1%). The main effect of neighborhood frequency was marginally significant, $F_s(1, 39) = 3.81$, $p = .05$, $MSE = 8.86$, and

---

[4] A close examination of the mean response latencies and error rates to each of the low-frequency words revealed that a few items had very high error rates and very long response latencies (e.g., *FUSE,* mean error rate = 27.5%, mean response latency = 732 ms). We suspected that these words were responsible for the item analysis being only marginally significant, through inflating the variance in the low-frequency word conditions and thereby reducing the statistical power of the test. To confirm this, we removed the slowest and most error-prone item from each of the low-frequency conditions (leaving 25 items in each condition) and recalculated the items ANOVA. In this analysis, both the main effect of neighborhood size and the main effect of neighborhood frequency were statistically significant, $F_i(1, 96) = 4.11$, $p < .05$, $MSE = 1,806.41$, and $F_i(1, 96) = 4.75$, $p < .05$, $MSE = 1,806.41$. Removing the same words from the item analyses of the low-frequency words in Experiment 1A and 1B had a similar effect, namely, lowering the p values for the neighborhood size and neighborhood frequency effects.

$F_i(1, 200) = 1.10$, $p > .10$, $MSE = 19.90$. Participants made fewer errors to words with higher frequency neighbors than to words with no higher frequency neighbors (2.7% vs. 3.3%). The main effect of neighborhood size was not significant (both $F\text{s} < 1$), nor were any of the interactions (all $p\text{s} > .20$).

Separate analyses of the high-frequency and low-frequency words were also conducted. For the high-frequency words, more errors were made to words with large neighborhoods than to words with small neighborhoods (2.4% vs. 1.5%, respectively), $F_s(1, 39) = 3.68$, $p = .06$, $MSE = 8.15$, and $F_i(1, 100) = 2.22$, $p > .10$, $MSE = 8.76$, consistent with the inhibitory neighborhood size effect witnessed in the response latency data. For the low-frequency words, there was no effect of neighborhood size (both $F\text{s} < 1$).

Summary of Experiments 1A–1D. The nonwords used in Experiment 1A had no neighbors, and thus, they should have generated very little lexical activity. Under these conditions, the multiple read-out model predicts that the degree of global lexical activity generated by a letter string will be a reliable cue as to the lexicality of that item. That is, words will produce significantly more lexical activation than nonwords, which would allow participants to use the Σ criterion for responding (i.e., participants will make most of their lexical decisions prior to lexical selection). Because words with large neighborhoods and words with higher frequency neighbors will produce more lexical activity than words with small neighborhoods and words without higher frequency neighbors, the model predicts that for low-frequency words, there should be a facilitatory neighborhood size effect and a facilitatory neighborhood frequency effect. This is essentially the pattern of results that was observed in this experiment.

Two additional results are of interest. First, the neighborhood size effect did not interact with word frequency, as both the high-frequency words and the low-frequency words exhibited a facilitatory neighborhood size effect. This result is perhaps not too surprising, because if most of the responses were based on the Σ criterion, then words with large neighborhoods would exceed the Σ criterion before words with small neighborhoods, regardless of the word's frequency. Second, the neighborhood frequency effect was modulated by word frequency. That is, responses to low-frequency words with higher frequency neighbors were faster and less error prone than responses to low-frequency words without higher frequency neighbors, whereas there were no such differences for the high-frequency words. This result suggests that for high-frequency words, the number of neighbors has a larger effect on lexical activity than the existence of higher frequency neighbors.

All of the nonwords used in Experiment 1B had small neighborhoods. Because at least the words with large neighborhoods would generate more lexical activity than the nonwords, responses to these words would have often been based on the Σ criterion. Thus, the multiple read-out model predicts that a facilitatory neighborhood size effect should be observed. However, because the words with small neighborhoods and the nonwords with small neighborhoods could not be reliably distinguished via the Σ criterion, the model also assumes that responses to small neighborhood words had to be based on the M criterion. As such, for those words, an inhibitory neighborhood frequency effect should have been observed. The end result should have been an interaction between neighborhood size and neighborhood frequency, with only the words with small neighborhoods exhibiting an inhibitory neighborhood frequency effect.

For the low-frequency words, words with large neighborhoods were responded to faster than words with small neighborhoods, consistent with the model's prediction. (For the high-frequency words, there was no effect of neighborhood size.) But the predicted interaction between neighborhood size and neighborhood frequency did not occur, as the neighborhood frequency effect for the small neighborhood words was not inhibitory. Instead, and contrary to the model's predictions, the neighborhood frequency effects were facilitatory and essentially equivalent for the small and large neighborhood words.

Two additional results are of note. First, both overall latencies and the word frequency effect were larger in this experiment than in Experiment 1A (41 ms vs. 26 ms), as indicated by a significant main effect of experiment, $F_s(1, 78) = 28.24$, $p < .001$, $MSE = 13{,}046.11$, and $F_i(1, 400) = 178.45$, $p < .001$, $MSE = 1{,}398.34$, and a significant interaction between experiment and word frequency, $F_s(1, 78) = 11.15$, $p < .01$, $MSE = 809.51$, and $F_i(1, 400) = 5.37$, $p < .05$, $MSE = 1{,}398.34$, in a combined analysis. The experiment effect confirms that the word–nonword discriminations were more difficult when the nonwords had small neighborhoods than when they had no neighbors. The interaction supports the idea that the M criterion was used more frequently in Experiment 1B than in Experiment 1A, because, according to the model, only the lexical selection process itself is sensitive to word frequency.

Second, unlike the results in Experiment 1A, for which a facilitatory neighborhood size effect was observed for low-frequency words and high-frequency words, in this experiment, the neighborhood size effect was modulated by word frequency. That is, for low-frequency words, the neighborhood size effect was facilitatory, whereas for high-frequency words, neighborhood size had little effect on response latencies or errors. Thus, as noted above, increasing the difficulty of the word–nonword discrimination not only increased the magnitude of the word frequency effect, it also eliminated the neighborhood size effect for high-frequency words. This result is consistent with the idea that the neighborhood size effect for high-frequency words in Experiment 1A was not a lexical selection effect (as it was eliminated in conjunction with an increase in the word frequency effect), but was instead a consequence of the extensive use of the Σ criterion for responding to high- as well as to low-frequency words.

In Experiment 1C, the nonwords were matched to the words on neighborhood size. As noted, under these conditions, it should not be possible for participants to use global lexical activation as a reliable cue for responding. According to the multiple read-out model, the majority of responses should therefore be made using the M criterion, and the Σ criterion should play very little (if any) role. Consequently, the model predicts a large inhibitory neighborhood frequency effect, and no effect of neighborhood size. The results are clearly at odds with both of these predictions, as the effect of neighborhood size and the effect of neighborhood frequency were facilitatory for low-frequency words. (There was no effect of neighborhood size nor of neighborhood frequency for the high-frequency words.)

Note that response latencies were slower in this experiment relative to those in Experiment 1A, $F_s(1, 78) = 40.66$, $p < .001$, $MSE = 22{,}385.35$, and $F_i(1, 400) = 430.70$, $p < .001$, $MSE =$

1,399.04, and relative to those in Experiment 1B, $F_s(1, 78) = 6.04$, $p < .05$, $MSE = 19{,}951.31$, and $F_i(1, 400) = 42.71$, $p < .001$, $MSE = 1{,}792.79$. This suggests that the word–nonword discriminations were more difficult in this experiment than in Experiments 1A and 1B. Also note that the word frequency effect in this experiment (53 ms) was larger than that observed in Experiment 1A (26 ms), $F_s(1, 78) = 33.99$, $p < .001$, $MSE = 905.60$, and $F_i(1, 400) = 15.28$, $p < .001$, $MSE = 1{,}399.04$, and in Experiment 1B (41 ms), $F_s(1, 78) = 6.97$, $p < .05$, $MSE = 928.98$, and $F_i(1, 400) = 1.98$, $p > .10$, $MSE = 1{,}792.79$. These results suggest that the M criterion was used more frequently in this experiment, as would be predicted by the multiple read-out model.

All of the nonwords used in Experiment 1D had large neighborhoods. As noted, according to the multiple read-out model, the presence of large neighborhood nonwords should make the word–nonword discriminations difficult enough to cause the Σ criterion to generally be set quite high. Thus, the model predicts that virtually all of the responses should be based on the M criterion, and an inhibitory neighborhood frequency effect should be observed. The model also predicts essentially a null neighborhood size effect because the Σ criterion would be used for responding only very occasionally, if at all. Recall that these were the same predictions the model made for Experiment 1C.

In this experiment, there was no effect of neighborhood size for the low-frequency words, as predicted by the model.[5] There was, however, no evidence of an inhibitory neighborhood frequency effect. In fact, the opposite was true, as words with higher frequency neighbors were responded to more rapidly than words without higher frequency neighbors.

There was some evidence of an inhibitory neighborhood size effect for the high-frequency words in this experiment. That is, response latencies and error rates to high-frequency words with large neighborhoods were slower and more error prone than responses to high-frequency words with small neighborhoods. This result was unexpected, as neighborhood size effects for high-frequency words are not common in the literature, and when observed, they have always been facilitatory (e.g., Sears et al., 1995, Experiment 1).

The multiple read-out model seems to be incapable of accounting for this particular result, because neighborhood size effects should be either facilitatory or nonexistent according to the model. However, it seems possible to explain this result in terms of the decision demands of a lexical decision task. More specifically, the increases in the overall response latencies for both the words and the nonwords across Experiments 1A–1D (Table 6) are consistent with an increase in the difficulty of the decision demands of the task because of the nature of the nonword context. In Experiment 1D, the nonwords were the most wordlike, as they had more neighbors, on average, than the words, and this produced the slowest response latencies to both high-frequency and low-frequency words. Because of the increased difficulty of distinguishing the words from the nonwords with large neighborhoods (relative to Experiment 1C), participants may have adopted a more cautious decision criterion for words with many neighbors, perhaps by invoking an additional checking process for items that generate very high levels of lexical activity. This would have allowed participants to keep their nonword error rate fairly low (such that it was similar to that of Experiment 1C), but it would have eliminated the advantage that usually occurs for low-frequency words with large neighborhoods. It would have also led to inhibition for large-neighborhood high-frequency words, which usually do not differ from high-frequency words with few neighbors. Thus, these results could be explained by assuming that under conditions involving highly wordlike nonwords, to keep their error rates low, participants adopted a conservative decision criterion for words with many neighbors, which eliminated the neighborhood size effect for low-frequency words and produced an inhibitory neighborhood size effect for high-frequency words.[6, 7]

Finally, note that response latencies were slower in this experiment relative to those in Experiment 1A, $F_s(1, 78) = 56.84$, $p < .001$, $MSE = 22{,}673.16$, and $F_i(1, 400) = 561.77$, $p < .001$, $MSE = 1{,}531.44$, and those in Experiment 1B, $F_s(1, 78) = 13.79$, $p < .001$, $MSE = 20{,}239.12$, and $F_i(1, 400) = 95.15$, $p < .001$,

---

[5] This finding is inconsistent with that of Sears et al.'s (1995) Experiment 5, in which there was a facilitatory neighborhood size effect under similar nonword conditions. The most likely reason for the difference between the results of Sears et al.'s Experiment 5 and those of Experiment 1D is that error rates were larger in Sears et al.'s experiment (6.6% for words and 9.0% for nonwords, vs. 4.0% for the low-frequency words and 4.5% for nonwords in Experiment 1D). This would suggest that Sears et al.'s participants were making more responses based on global lexical activity than the participants in Experiment 1D (see Grainger & Jacobs, 1996, Experiment 1D). (Note that Grainger & Jacobs, 1996, suggested this possibility in their discussion of Sears et al.'s [1995] Experiment 5.)

To address this possibility, we used the same word and nonword stimuli in an experiment in which instructions emphasizing speed over accuracy were provided to a new group of participants. If Sears et al.'s results were due to their participants making many of their responses on the basis of global lexical activity rather than lexical selection, then inducing the same type of responding through instructions should have produced a facilitatory neighborhood size effect. This was in fact the case. There was also a significant Word Frequency × Neighborhood Size interaction, reflecting that large neighborhoods facilitated responses to only low-frequency words (i.e., there was no effect of neighborhood size for the high-frequency words). Specifically, for the low-frequency words, words with large neighborhoods were responded to an average of 12 ms faster than words with small neighborhoods, an effect size similar to the 14-ms effect reported by Sears et al. In addition, the error rates to the low-frequency words (7.1%) were similar to those observed in Sears et al.'s Experiment 5 (6.6%), as were the error rates to the nonwords (8.4% here vs. 9.0% in Sears et al.'s experiment).

[6] We thank an anonymous reviewer for suggesting this possibility.

[7] We were able to replicate both of these findings in two experiments designed specifically for this purpose. In one of these experiments, 40 participants responded to high-frequency words with large neighborhoods and to high-frequency words with no neighbors; all of the nonwords had large neighborhoods. The words with large neighborhoods were responded to an average of 13 ms more slowly than the words with no neighbors, $F_s(1, 39) = 6.16$, $p < .05$, $MSE = 541.46$, and $F_i(1, 38) = 1.22$, $p > .10$, $MSE = 1{,}666.96$, and more errors were made to the words with large neighborhoods than to the words with no neighbors, $F_s(1, 39) = 4.04$, $p = .05$, $MSE = 13.07$, and $F_i(1, 38) = 1.02$, $p > .10$, $MSE = 25.78$. This result thus replicated the inhibitory neighborhood size effect witnessed for high-frequency words in Experiment 1D. In the other experiment, a different group of 40 participants responded to low-frequency words with large neighborhoods and to low-frequency words with no neighbors. All of the nonwords had large neighborhoods. In this experiment, there was no effect of neighborhood size in the analysis of response latencies (both Fs < 1), or in the analysis of error rates, $F_s(1, 39) = 2.18$, $p > .10$, $MSE = 9.17$, and $F_i < 1$. This result thus replicated the null effect of neighborhood size witnessed for low-frequency words in Experiment 1D.

MSE = 1,925.19, and were similar to those in Experiment 1C, $F_s(1, 78) = 1.11$, $p > .10$, *MSE* = 29,578.36, and $F_i(1, 400) = 11.88$, $p < .01$, *MSE* = 1,925.88. This suggests that the word–nonword discriminations were more difficult in this experiment than in Experiments 1A and 1B and were of similar difficulty to those in Experiment 1C (see Table 6). Also note that the word frequency effect was larger in this experiment (46 ms) than in Experiment 1A (26 ms), $F_s(1, 78) = 18.65$, $p < .001$, *MSE* = 877.69, and $F_i(1, 400) = 8.45$, $p < .01$, *MSE* = 1,531.44, and was similar to the word frequency effects observed in Experiment 1B (41 ms), $F_s(1, 78) = 1.21$, $p > .10$, *MSE* = 901.08, and $F_i < 1$, and in Experiment 1C (53 ms), $F_s(1, 78) = 2.26$, $p > .10$, *MSE* = 997.16, and $F_i < 1$ (see Table 7).[8]

Simulations with the multiple read-out model. We conducted several simulations using the multiple read-out model to examine the model's ability to account for the facilitatory neighborhood size and facilitatory neighborhood frequency effects observed in Experiment 1.[9] In these simulations, the same parameters adopted by Grainger and Jacobs (1996) were used, including setting the activation threshold (the M criterion) to 0.67. The four-letter and five-letter lexicons used in the simulations consisted of words with Kucera and Francis (1967) frequencies greater than zero. The four-letter lexicon consisted of 1,580 words; the five-letter lexicon consisted of 2,124 words.

Word identification latencies were simulated by the number of processing cycles required for a word's lexical unit to reach the M criterion, or the number of cycles elapsed when the summed lexical activation generated by the word exceeded the Σ criterion. Unlike Grainger and Jacobs's (1996) simulations, our simulations were not stochastic, in that the M, Σ, and T criteria did not vary randomly around a mean value. Grainger and Jacobs (1996) used stochastics in their simulations to create variability in the simulated response latencies and error rates to each item. Our goal was to ascertain whether the model could capture the general pattern of effects witnessed in our experiments, and, therefore, stochastics were not necessary. Note that this should be of no consequence in the interpretation of the results of our simulations.

The purpose of our first simulation was to determine whether words with higher frequency neighbors would require more processing cycles to reach the M criterion than words without higher frequency neighbors while, at the same time, there would be no difference between words with large versus small neighborhoods. Consequently, in this simulation, the Σ and T criteria were not used (note that this makes the multiple read-out model's predictions identical to those of the interactive-activation model). According to the model, when only the M criterion is used for responding, participants must wait until lexical selection is completed before making a response. In this situation, words with higher frequency neighbors should be responded to more slowly than words without higher frequency neighbors, and there should be no evidence of a neighborhood size effect. The mean number of processing cycles required to reach the M criterion for the words used in Experiment 1 is shown in Table 8.

Table 8
Mean Number of Processing Cycles and Mean Summed Lexical Activity for the Word Stimuli Used in Experiments 1A–1D

Neighborhood frequency	Neighborhood size: Small	Neighborhood size: Large
Low-frequency words		
No HF N	17.18 (.39)	17.17 (.42)
HF N	17.82 (.42)	17.80 (.45)
High-frequency words		
No HF N	16.94 (.40)	16.90 (.45)
HF N	17.36 (.42)	17.37 (.47)

Note. HF N = higher frequency neighbors. Mean summed lexical activity appears in parentheses.

These data were submitted to a 2 (word frequency: high or low) × 2 (neighborhood size: small or large) × 2 (neighborhood frequency: no higher frequency neighbors or higher frequency neighbors) factorial ANOVA. There was a main effect of word frequency, $F_i(1, 200) = 33.90$, $p < .001$, *MSE* = 0.19, as high-frequency words required an average of 0.35 fewer processing cycles than low-frequency words (17.14 vs. 17.49). The main effect of neighborhood frequency was also significant, $F_i(1, 200) = 79.42$, $p < .001$, *MSE* = 0.19. Words with higher frequency neighbors required an average of 0.54 more processing cycles than words without higher frequency neighbors (17.58 vs. 17.04).

The main effect of neighborhood size was not significant ($F_i < 1$). The mean number of processing cycles for the words with small neighborhoods was 17.32; for the words with large neighborhoods, it was 17.31. The only other effect that approached statistical significance was the interaction between word frequency and neighborhood frequency, $F_i(1, 200) = 2.43$, $p = .12$, *MSE* = 0.19. An examination of Table 8 suggests that the predicted inhibitory effect of higher frequency neighbors was slightly larger for the low-frequency words (.63) than for the high-frequency words (.44). The results of this simulation are thus very clear—when only the M criterion is used for responding, the multiple read-out model predicts that for these stimuli, there should be a word frequency effect and an inhibitory neighborhood frequency effect, but no neighborhood size effect.

In our second simulation, we sought to determine whether the model could account for the facilitatory neighborhood size and facilitatory neighborhood frequency effects witnessed in Experiment 1C, for which the nonwords had either a large or a small neighborhood. Recall that in this situation, according to our reasoning, because the degree of lexical activity generated by the words and the nonwords would be very similar, it should not be

---

[8] Although the neighborhood size and the neighborhood frequency effects were often only marginally significant in the item analyses of the individual experiments, these effects were always significant in combined analyses of the experiments. Specifically, the neighborhood size effect was significant in an item analysis when the stimuli from Experiments 1A and 1B were combined, $F(1, 400) = 5.51$, $p < .05$, *MSE* = 1,398.34, when the stimuli from Experiments 1A and 1C were combined, $F(1, 400) = 6.54$, $p < .05$, *MSE* = 1,399.04, and when the stimuli from Experiments 1B and 1C were combined, $F(1, 400) = 5.22$, $p < .05$, *MSE* = 1,792.79. The neighborhood frequency effect was significant in the same analyses, $F(1, 400) = 5.38$, $p < .05$, *MSE* = 1,398.34; $F(1, 400) = 6.34$, $p < .05$, *MSE* = 1,399.04; and $F(1, 400) = 4.23$, $p < .05$, *MSE* = 1,792.79, respectively.

[9] We thank Walter van Heuven for providing us with the software we used to implement the multiple read-out model.

possible to reliably distinguish the words from the nonwords on the basis of global lexical activation. We therefore assumed that most (if not all) of the responses would have been made using the M criterion, which should have produced an inhibitory neighborhood frequency effect and a null neighborhood size effect (as in the first simulation). Nonetheless, facilitatory neighborhood size and facilitatory neighborhood frequency effects were observed in this experiment. Because this is the most problematic set of results for the model, we focused our simulation efforts (and their presentation herein) on this experiment.

According to the model, the only way facilitatory neighborhood effects could have occurred in this experiment was if (a) the words with large neighborhoods produced more lexical activation than the words with small neighborhoods, (b) the words with higher frequency neighbors produced more lexical activation than the words without higher frequency neighbors, and (c) participants relied on the Σ criterion rather than the M criterion when responding to words with large neighborhoods and to words with higher frequency neighbors. In this simulation, we therefore incorporated the Σ criterion (and the T criterion) and attempted to produce the pattern of effects observed in Experiment 1C.

To do so, we first had to determine the degree of lexical activation generated by the word and the nonword stimuli used in Experiment 1C.[10] The word data are presented in Table 8 (the mean summed lexical activation produced by the nonwords with small neighborhoods was .30; for the nonwords with large neighborhoods it was .41). As can be seen in the table, words with large neighborhoods produced more lexical activation than words with small neighborhoods, $F_{\mathrm{i}}(1, 200) = 139.06$, $p < .001$, $MSE = 0.001$, and words with higher frequency neighbors produced more lexical activation than words without higher frequency neighbors, $F_{\mathrm{i}}(1, 200) = 46.41$, $p < .001$, $MSE = 0.001$. This was true for both the high-frequency words and the low-frequency words. That is, a separate analysis of the high-frequency words revealed an effect of neighborhood size and an effect of neighborhood frequency, $F_{\mathrm{i}}(1, 100) = 82.76$, $p < .001$, $MSE = 0.001$, and $F_{\mathrm{i}}(1, 100) = 14.03$, $p < .001$, $MSE = 0.001$, respectively, as did an analysis of the low-frequency words, $F_{\mathrm{i}}(1, 100) = 56.46$, $p < .001$, $MSE = 0.001$, and $F_{\mathrm{i}}(1, 100) = 36.55$, $p < .001$, $MSE = 0.001$, respectively.

There was also an effect of word frequency, $F_{\mathrm{i}}(1, 200) = 10.28$, $p < .01$, $MSE = 0.001$, and a marginally significant interaction between word frequency and neighborhood size, $F_{\mathrm{i}}(1, 200) = 3.36$, $p = .06$, $MSE = 0.001$. High-frequency words produced more lexical activation than low-frequency words, and the difference between the words with large neighborhoods and the words with small neighborhoods was slightly larger for the high-frequency words. Together these analyses suggest that the conditions necessary for use of the Σ criterion to produce facilitatory neighborhood effects were present—words with large neighborhoods and words with higher frequency neighbors produced more lexical activation than words with small neighborhoods and words without higher frequency neighbors.

To simulate the facilitatory neighborhood effects obtained in Experiment 1C, we kept the M criterion at 0.67 and manipulated the values of the Σ and T criteria to produce facilitatory neighborhood size and neighborhood frequency effects for the low-frequency words.[11] As can be seen in Table 9, this was possible, but not without producing the same effects for the high-frequency words (note that in the experiment, there were no neighborhood effects for high-frequency words). There was a significant effect of neighborhood size and of neighborhood frequency, $F_{\mathrm{i}}(1, 193) = 13.39$, $p < .001$, $MSE = 1.04$, and $F_{\mathrm{i}}(1, 193) = 7.62$, $p < .01$, $MSE = 1.04$, but neither of these effects interacted with word frequency (both *p*s $> .30$). This outcome makes sense because, as noted above, high-frequency words with large neighborhoods and higher frequency neighbors produced more lexical activation than high-frequency words with

---

[10] Grainger and Jacobs (1996) used the summed lexical activity generated by a stimulus after seven cycles of processing as their index of lexical activation (denoted as σ in the model). Grainger and Jacobs noted that after seven cycles of processing, five-letter words typically produce more summed lexical activity than four-letter words. Because σ is used to set the Σ and T criteria on each trial (e.g., if $\sigma > .25$, then T = 20 cycles; otherwise, T = 18 cycles), large differences in the lexical activity produced by four- and five-letter words are problematic, because that would necessitate the adoption of different σ thresholds for four- and for five-letter words (e.g., for four-letter stimuli, if $\sigma > .25$, then T = 20 cycles; otherwise, T = 18 cycles; for five-letter stimuli, if $\sigma > .35$, then T = 20 cycles; otherwise, T = 18 cycles). To keep the σ values generated by four- and five-letter stimuli similar so that the same σ threshold could be used, Grainger and Jacobs chose to reduce the lexical activity generated by the five-letter stimuli. This was accomplished by reducing the letter-to-word excitation parameter from 0.07 to 0.06 in their simulations involving five-letter stimuli (see Grainger & Jacobs, 1996, for their rationale). With this modification, Grainger and Jacobs reported that the σ values generated by the five-letter stimuli were within the same range as those generated by the four-letter stimuli.

With our stimuli, however, this modification did not appreciably reduce the difference between the four- and the five-letter stimuli. Consequently, in addition to reducing the letter-to-word excitation parameter for the five-letter stimuli, for the four-letter stimuli, we decided to use the σ values generated after eight cycles of processing (which are larger than the σ values generated after seven cycles of processing). This statistically equated the four- and five-letter words on σ. Note that it did not, of course, eliminate the difference in the lexical activation generated by the four- and five-letter stimuli words during subsequent processing cycles. As noted below (see Footnote 11), this necessitated using slightly different Σ criteria for the four- and five-letter stimuli in our simulations.

[11] In the first simulation of Experiment 1C, the T criterion was set using the following rule: If the summed lexical activity generated by a stimulus (σ) was greater than .37 then T = 20; otherwise, T = 17. For the Σ criterion, slightly different rules were required for the four- and five-letter stimuli, because the five-letter stimuli produced more lexical activation throughout their processing (as previously noted). For the four-letter stimuli, if the summed lexical activity generated was greater than .43, then the Σ criterion was set at .72; otherwise, $\Sigma = 1.5$. For the five-letter words, if the summed lexical activity was greater than .43, then $\Sigma = .84$; otherwise, $\Sigma = 1.5$. In the second simulation of Experiment 1C, for the T criterion, if the summed lexical activity was greater than .37, then T = 20; otherwise, T = 17. For the four-letter stimuli, if the summed lexical activity generated was greater than .49, then $\Sigma = .72$; otherwise, $\Sigma = 1.5$. For the five-letter words, if the summed lexical activity was greater than .49, then $\Sigma = .84$; otherwise, $\Sigma = 1.5$.

small neighborhoods and no higher frequency neighbors; therefore, they too would be affected by use of the Σ criterion.[12]

A much more serious problem was that the predicted error rate to the nonwords with large neighborhoods was 20.7% in this simulation, whereas in the experiment itself, the error rate to these stimuli was 5.4%. The problem here stems from many of these nonwords generating as much or even more lexical activity than the words (mean summed lexical activity for the large neighborhood nonwords = .41, range = .33–.59; for the words, mean summed lexical activity = .43, range = .33–.56). Consequently, when many of the *word* responses were based on the Σ criterion during the simulation, many of the large neighborhood nonwords also exceed the criterion and hence resulted in errors.

Given the seriousness of this problem, we conducted another simulation in which we chose values of the Σ and T criteria that produced a large neighborhood nonword error rate similar to that observed in the experiment (5.7%). When the nonword error rate was constrained in this manner, the facilitatory neighborhood size effect disappeared and the neighborhood frequency effect was once again inhibitory (with the exception of the high-frequency words with large neighborhoods). These data are shown in Table 10.

In summary, our attempts to simulate the results of Experiment 1C were not successful. As we predicted, when the nonwords are matched to the words on neighborhood size, the degree of lexical activation is not useful for distinguishing the words from the nonwords. This prohibits the extensive use of the Σ criterion necessary to produce facilitatory neighborhood effects.[13]

Discussion

Together, the results of these experiments seriously challenge the multiple read-out model's most basic assumptions. First, recall that the model assumes that facilitatory effects of neighborhood size will occur only when participants rely on the Σ criterion for responding. Further, the extent to which the Σ criterion is used will depend critically on the extent to which the words and nonwords can be distinguished from one another on the basis of the lexical activation they generate. More specifically, according to the model, when the nonwords have no neighbors or when they have small neighborhoods, they will be relatively easy to distinguish from the words on the basis of lexical activation, and thus, a

Table 9
Simulation of Facilitatory Neighborhood Effects in Experiment 1C: Mean Number of Processing Cycles

	Neighborhood size	
Neighborhood frequency	Small	Large
Low-frequency words		
No HF N	17.11 (15.3)	16.91 (0.0)
HF N	16.97 (3.8)	16.41 (0.0)
High-frequency words		
No HF N	16.94 (0.0)	16.68 (0.0)
HF N	16.88 (7.6)	15.79 (0.0)

Note. HF N = higher frequency neighbors. Simulated error percentages appear in parentheses.

Table 10
Simulation of Experiment 1C: Mean Number of Processing Cycles

	Neighborhood size	
Neighborhood frequency	Small	Large
Low-frequency words		
No HF N	17.18 (15.3)	17.17 (0.0)
HF N	17.83 (3.8)	17.54 (0.0)
High-frequency words		
No HF N	16.94 (0.0)	16.90 (0.0)
HF N	17.36 (7.6)	16.44 (0.0)

Note. HF N = higher frequency neighbors. Simulated error percentages appear in parentheses.

---

[12] Attempts to eliminate the neighborhood effects for the high-frequency words by manipulating the Σ and T criteria resulted in the effects being eliminated for the low-frequency words as well. As a consequence, we could not simulate the interaction between word frequency and neighborhood size observed in this experiment and in other studies (i.e., a facilitatory neighborhood size effect for low-frequency words but not for high-frequency words; e.g., Andrews, 1989, 1992). Note that the model can simulate this interaction under some circumstances, as Grainger and Jacobs (1996) demonstrated in their simulations of Andrews's (1989, 1992) lexical decision experiments. Unlike our stimuli, however, Andrews's high-frequency words with large neighborhoods did not produce much more lexical activation than her high-frequency words with small neighborhoods (.43 vs. .42, respectively, for Andrews's 1989 stimuli, and .45 vs. .43, respectively, for Andrews's 1992 stimuli, as compared with .46 vs. .41 for our stimuli). As a result, few of Andrews's high-frequency words with large neighborhoods would have been affected by use of the Σ criterion, which would allow the model to simulate a null effect of neighborhood size for high-frequency words.

[13] The mean summed lexical activity generated by the nonwords with no neighbors used in Experiment 1A was .14, with a range of .08 to .24. In our simulations of Experiment 1A, we could produce facilitatory neighborhood size and facilitatory neighborhood frequency effects both for the high-frequency and the low-frequency words (as observed in the experiment). The mean summed lexical activity generated by the nonwords with small neighborhoods used in Experiment 1B was .30, with a range of .09 to .42. Our simulations of this experiment met with varying degrees of success. We could simulate facilitatory neighborhood size and neighborhood frequency effects for the low-frequency words, but not without producing the same effects for high-frequency words. Conversely, we could simulate null effects of neighborhood size and of neighborhood frequency for the high-frequency words, but not without eliminating these effects for the low-frequency words as well. Finally, the mean summed lexical activity generated by the nonwords with large neighborhoods used in Experiment 1D was .40, with a range of .25 to .59. When the simulated nonword error rate was constrained to be similar to that observed in the experiment (i.e., 4.8% in the simulation vs. 4.9% in the experiment), for the low-frequency words, there was an inhibitory effect of neighborhood frequency and no effect of neighborhood size. For the high-frequency words, there was an interaction between neighborhood size and neighborhood frequency, such that the words with higher frequency neighbors exhibited a facilitatory neighborhood size effect, whereas the words without higher frequency neighbors did not.

facilitatory neighborhood size effect should occur. Conversely, when the nonwords have large neighborhoods or when they have the same neighborhood sizes as the words, the Σ criterion will not be used for responding and no neighborhood size effects should occur.

In Experiment 1A, the nonwords had no orthographic neighbors, and in Experiment 1B, the nonwords had small neighborhoods, and thus the model predicted that a facilitatory neighborhood size effect should have been observed in both of these experiments. This was in fact the case. Similarly, in Experiment 1D, the nonwords had large neighborhoods, and the model's prediction of a null neighborhood size effect for the low-frequency words was upheld.

The problem for the model in terms of the neighborhood size effect was the facilitatory effect observed in Experiment 1C, in which the words and the nonwords were matched on neighborhood size (and hence could not be distinguished from one another on the basis of lexical activation). The model clearly predicts that in this situation, a neighborhood size effect should not occur; yet in this experiment, the largest facilitatory neighborhood size effect of all of the experiments was observed. Indeed, the model predicts that for low-frequency words, the largest facilitatory neighborhood size effect should have been observed in Experiment 1A (nonwords with no neighbors), but the neighborhood size effect in that experiment was smaller than that observed in Experiment 1C (10.5 ms vs. 19.5 ms, respectively), $t(78) = 1.62$, $p = .05$, one-tailed.

In addition to this problem, the predicted relation between the word frequency effect and the neighborhood size effect was not observed. Recall that, according to the model, the neighborhood size effect and the word frequency effect should be negatively correlated, with larger neighborhood size effects corresponding to smaller word frequency effects (because of increased use of the Σ criterion) and smaller neighborhood size effects corresponding to larger word frequency effects (because of increased use of the M criterion). In Experiment 1A, the neighborhood size effect was 10.5 ms and the word frequency effect was 26 ms, and in Experiment 1C, the neighborhood size effect was 19.5 ms, and the word frequency effect was 53 ms. Thus, contrary to the prediction of the model, a larger neighborhood size effect was associated with a larger word frequency effect, not a smaller one.[14] Overall, then, the model does not appear to provide a particularly good account of the neighborhood size effects observed in these experiments.

With respect to the neighborhood frequency effect, the data are even more problematic for the model. Because the multiple read-out model assumes that inhibitory neighborhood frequency effects are due to the use of the M criterion, inhibitory neighborhood frequency effects should be observed whenever the words cannot be reliably distinguished from the nonwords on the basis of lexical activation (i.e., when the nonwords have large neighborhoods or when the words and the nonwords are matched on neighborhood size). In Experiment 1C, the words and the nonwords were matched on neighborhood size, and in Experiment 1D, the nonwords had large neighborhoods; yet in both of these experiments, facilitatory neighborhood frequency effects were observed. Indeed, in all of the experiments, the effect of higher frequency neighbors was facilitatory.[15] Although these results are quite consistent with those reported by Andrews (1989, 1992), Forster and Shen (1996), and Sears et al. (1995), they are completely opposite to what the multiple read-out model predicts.

Experiment 2

The most problematic outcome for the multiple read-out model in the previous experiments was the absence of an inhibitory neighborhood frequency effect, particularly in Experiments 1C and 1D, when responses should have been primarily based on the M criterion. Instead, the effect of higher frequency neighbors was facilitatory, not inhibitory, a result that is very difficult for the model to accommodate.

Because these results are so problematic for the model, we felt it was necessary to determine whether these effects would be replicated in a new set of experiments. Accordingly, in Experiment 2, the focus was solely on the neighborhood frequency effect, and again lexical decision conditions were created in which it should not be possible to rely on the Σ criterion for responding. More specifically, in Experiment 2A, all of the words and the nonwords had small neighborhoods, and the words had either no higher frequency neighbors or exactly one higher frequency neighbor. In Experiment 2B, all of the words and the nonwords had large neighborhoods, and the words had either no higher frequency neighbors or exactly one higher frequency neighbor. (The word stimuli in both experiments were of low frequency.) Because the words and the nonwords were matched on neighborhood size in each of these experiments, it should not be possible to reliably

---

[14] A more extreme comparison is possible when the nonwords are orthographically illegal letter strings. We conducted this experiment and found the word frequency effect to be 11 ms, $F_s(1, 39) = 19.17$, $p < .001$, $MSE = 559.46$, and the neighborhood size effect to be 6 ms, $F_s(1, 39) = 8.89$, $p < .01$, $MSE = 559.55$, which further reinforces this observation. Also relevant is an experiment by Lewellen et al. (1993, Experiment 2), in which nonword orthography (illegal vs. legal) was manipulated across two blocks of trials. In the orthographically illegal nonword block, the word frequency effect was 4 ms and the neighborhood size effect was 22 ms; in the orthographically legal nonword block, the word frequency effect was 21 ms and the neighborhood size effect was 34 ms.

[15] Carreiras et al. (1997) have suggested that failures to observe inhibitory neighborhood frequency effects in English may be due to not controlling for phonological inconsistency among a word's orthographic neighbors (e.g., using the target word *WARM*, which has an inconsistent higher frequency neighbor, *FARM*). To address this possibility, we conducted a post hoc analysis in which all the target words that had inconsistent neighbors were removed and the remaining stimuli were reanalyzed. (The removal of the stimuli with inconsistent neighbors did not appreciably change any of the stimulus characteristics of the remaining target words. That is, the target words were still closely matched for word frequency, neighborhood size, and the normative frequency of the highest frequency neighbor.) The important findings of this reanalysis are as follows. First, the facilitatory neighborhood size effects of Experiments 1A, 1B, and 1C were still statistically significant. Second, for the high-frequency words in Experiment 1D, the inhibitory neighborhood size effect was still statistically significant, and there was no effect of neighborhood size for the low-frequency words. The facilitatory neighborhood frequency effects observed in Experiments 1A, 1B, and 1D were attenuated or eliminated in this reanalysis (presumably because of loss of statistical power); however, in no case were they inhibitory. The facilitatory neighborhood frequency effect observed in Experiment 1C, however, remained statistically significant. Thus, there was no evidence to support Carreiras et al.'s (1997) suggestion that inhibitory neighborhood frequency effects would be observed in conditions for which the presence of phonologically inconsistent orthographic neighbors has been controlled.

distinguish the words from the nonwords on the basis of lexical activity. Consequently, according to the model, in these situations, participants should set their Σ criterion quite high and virtually all responses should be based on the M criterion, in which case the words with a higher frequency neighbor should be responded to more slowly than the words without a higher frequency neighbor. As in Experiment 1, the model's predictions and its ability to account for the experimental data were tested in several simulations.

Method

Participants. There were 80 undergraduate students from the University of Calgary who participated in the experiment; 40 participated in Experiment 2A (words and nonwords with small neighborhoods), and 40 participated in Experiment 2B (words and nonwords with large neighborhoods). All were native English speakers and reported that they had normal or corrected-to-normal vision. None of these individuals participated in any of the previous experiments, or in more than one of the present experiments.

Stimuli. In each experiment, there were 30 words in each of two stimulus conditions, for a total of 60 words. (The complete sets of experimental words used in Experiments 2A and 2B are presented in the Appendix, and the descriptive statistics for these stimuli are listed in Table 11.)

In each experiment, half of the words in each condition were four-letter words, and the other half of the words were five-letter words. In Experiment 2A, all of the words were of low frequency (mean Kucera and Francis [1967] normative frequency = 19.6, range = 1–49), and had small neighborhoods (range = 1–5 neighbors; mean neighborhood size = 3.4). In Experiment 2B, all of the words were of low frequency (mean Kucera and Francis [1967] normative frequency = 22.9, range = 1–48), and had large neighborhoods (range = 6–17 neighbors; mean neighborhood size = 9.0). Approximately two thirds of the words used in these experiments had been used in Experiments 1A–1D.

In each experiment, the single factor manipulated was neighborhood frequency: The words had either no neighbors of higher frequency or exactly one neighbor of higher frequency. For the words with a higher frequency neighbor, the mean Kucera and Francis (1967) normative frequency of the highest frequency neighbor of each word was 301.1 in Experiment 2A and 290.1 in Experiment 2B. For the words with no higher frequency neighbors, in each experiment, the mean frequency of the highest frequency neighbor of each word was substantially lower than the mean target frequency.

Table 11
Mean Word Frequency and Neighborhood Size (N) for the Word Stimuli Used in Experiments 2A (Words and Nonwords with Small Neighborhoods) and 2B (Words and Nonwords with Large Neighborhoods)

	Neighborhood frequency	
Stimulus characteristic	No HF N	One HF N
Experiment 2A		
Frequency	19.1	20.1
N	3.3	3.5
NF	8.0	301.0
Experiment 2B		
Frequency	23.8	23.1
N	8.7	9.4
NF	13.6	290.1

Note. HF N = higher frequency neighbor(s). NF refers to the average frequency of the highest frequency neighbor.

Table 12
Mean Response Latencies (in Milliseconds), Standard Errors, and Error Rates (Percentages) in Experiments 2A (Words and Nonwords with Small Neighborhoods) and 2B (Words and Nonwords with Large Neighborhoods)

	Neighborhood frequency	
Experiment	No HF N	One HF N
2A	561 (8.5, 4.2)	545 (7.7, 4.2)
2B	603 (11.4, 3.9)	588 (9.2, 3.8)

Note. HF N = higher frequency neighbor(s). Standard errors and error rates, respectively, appear in parentheses.

In each experiment, the nonword stimuli consisted of 30 four-letter and 30 five-letter orthographically legal and pronounceable letter strings. In Experiment 2A, all the nonwords had small neighborhoods (range = 1–5 neighbors), with a mean neighborhood size of 3.4. These nonwords were a subset of those used in Experiment 1B. In Experiment 2B, all the nonwords had large neighborhoods (range = 6–17 neighbors), with a mean neighborhood size of 8.9. These nonwords were a subset of those used in Experiment 1D.

Design. In each experiment, the two neighborhood frequency conditions (no higher frequency neighbors or one higher frequency neighbor) produced a one-factor repeated measures design. Response latencies and error rates were submitted to a one-factor repeated measures ANOVA. Both subject (F_s) and item (F_i) analyses were performed.

Apparatus and procedure. The apparatus was identical to that used in Experiments 1A–1D. In each experiment, participants completed 30 practice trials prior to the collection of data. In Experiment 2A, the practice trials consisted of 15 low-frequency words and 15 orthographically legal and pronounceable nonwords with small neighborhoods (i.e., at least one and no more than five neighbors). In Experiment 2B, the practice trials consisted of 15 low-frequency words and 15 orthographically legal and pronounceable nonwords with large neighborhoods (i.e., with at least six neighbors). (These practice stimuli were not used in the experiments, and the data from these practice trials were not analyzed.) Following the practice trials, the participants were provided with feedback as to the mean latency and accuracy (percentage of errors) of their responses, and during the experimental trials, this information was presented every 30 trials. Participants were instructed to respond as quickly as possible while keeping their error rate below 5%. The order in which the stimuli were presented in the experiments was randomized separately for each participant.

Results

The mean response latencies of correct responses and the mean error rates for each experiment are shown in Table 12. The mean response latency and percentage of errors for the small neighborhood nonwords used in Experiment 2A were 607 ms and 4.4%, respectively. The mean response latency and percentage of errors for the large neighborhood nonwords used in Experiment 2B were 695 ms and 6.0%, respectively.

Experiment 2A: Words and nonwords with small neighborhoods. In the analysis of the response latencies, the effect of neighborhood frequency was significant, $F_s(1, 39) = 13.60$, $p <$

.01, *MSE* = 371.11, and $F_i(1, 58) = 1.29$, $p > .10$, *MSE* = 2,605.07. Responses to words with one higher frequency neighbor were an average of 16 ms faster than responses to words without higher frequency neighbors. In the analysis of the error data, the effect of neighborhood frequency was not significant (both *F*s < 1).

Experiment 2B: Words and nonwords with large neighborhoods. In the analysis of the response latencies, the effect of neighborhood frequency was significant, $F_s(1, 39) = 10.52$, $p < .01$, *MSE* = 465.40, and $F_i(1, 58) = 1.09$, $p > .10$, *MSE* = 4,408.08. Responses to words with one higher frequency neighbor were an average of 15 ms faster than responses to words without higher frequency neighbors. In the analysis of the error data, the effect of neighborhood frequency was not significant (both *F*s < 1).

Simulations with the multiple read-out model. In our first set of simulations, we determined whether the words with one higher frequency neighbor would require more processing cycles to reach the M criterion than the words without higher frequency neighbors. For Experiment 2A, words with one higher frequency neighbor did require more processing cycles to reach the M criterion than did words without higher frequency neighbors (17.85 vs. 17.19), $F_i(1, 58) = 52.19$, $p < .001$, *MSE* = 0.13. This was true of the stimuli used in Experiment 2B as well, with words with one higher frequency neighbor requiring an average of 17.59 processing cycles to reach the M criterion and words without higher frequency neighbors requiring an average of 17.17 cycles, $F_i(1, 58) = 14.53$, $p < .001$, *MSE* = 0.19. Thus, as expected, if only the M criterion is used for responding, then the multiple read-out model predicts that there should be an inhibitory neighborhood frequency effect in both of these experiments.

In our second set of simulations, we incorporated the Σ and T criteria and attempted to produce the facilitatory neighborhood frequency effect observed in Experiments 2A and 2B. For Experiment 2A, words with one higher frequency neighbor produced more lexical activation than words without higher frequency neighbors (.41 vs. .39), $F_i(1, 58) = 11.71$, $p < .01$, *MSE* = 0.001, which, according to the model, could allow the Σ criterion to be used for responding (mean lexical activation produced by the nonwords = .29; range = .09–.42). Nonetheless, when choosing values of the Σ and the T criteria that produced a nonword error rate similar to that observed in the experiment (i.e., 6.6% in the simulation vs. 4.4% in the experiment), the best outcome we could obtain was a null effect of neighborhood frequency. Specifically, for the words with one higher frequency neighbor, the mean number of processing cycles was 16.83, and for the words without higher frequency neighbors, it was 17.19, $F_i(1, 55) = 2.78$, $p > .10$, *MSE* = 0.67. (The simulated error rates for the words were 6.6% in the no higher frequency neighbor condition and 3.3% in the one higher frequency neighbor condition.)

For Experiment 2B, words with one higher frequency neighbor produced more lexical activation than words without higher frequency neighbors (.44 vs. .42), $F_i(1, 58) = 9.74$, $p < .01$, *MSE* = 0.001 (mean lexical activation produced by the nonwords = .40; range = .36–.51). But again, when the nonword error rate in the simulation was constrained to be similar to that observed in the experiment (6.6% in the simulation vs. 6.0% in the experiment), the best outcome was a null effect of neighborhood frequency. Specifically, for the words with one higher frequency neighbor, the mean number of processing cycles was 17.04, and for the words without higher frequency neighbors, it was 17.16 ($F_i < 1$). (The simulated error rates for the words were 3.3% in both conditions.) It was possible to simulate a facilitatory neighborhood frequency effect, but not without simulating a nonword error rate much larger than that observed in the experiment. In this simulation, the mean number of processing cycles for the words with one higher frequency neighbor was 16.05, and for the words without higher frequency neighbors, it was 16.77, $F_i(1, 56) = 7.89$, $p < .01$, *MSE* = 0.95. However, the simulated nonword error rate was 21.6%.

Two other points about these simulations should be noted. As demonstrated in the first set of simulations, if only the M criterion is used, then an inhibitory neighborhood frequency effect is predicted in both experiments. To eliminate this effect, as we were able to do in our simulations, many of the words with one higher frequency neighbor have to be responded to with the Σ criterion in the simulations. But an examination of the individual items responded to with the Σ criterion reveals two difficulties with this procedure. The first is that there is a tendency for the words that require the greatest number of processing cycles to reach the M criterion to also require the fewest number of processing cycles to reach the Σ criterion (this was also the case in the simulations of Experiment 1C). In effect, then, according to these simulations, the most difficult words (in terms of the time required for lexical selection) should be responded to as quickly or even more quickly than the easiest words—a very counterintuitive notion. One way to illustrate this is to correlate the number of processing cycles required to reach the M criterion with the number of processing cycles required to reach the Σ criterion. In the one higher frequency neighbor condition of Experiment 2A, this correlation was −.53 ($p < .05$), and in Experiment 2B, it was −.78 ($p < .05$). Second, and related, is that there was no evidence for such a relation within the item response latencies. In Experiment 2A, the correlation between the mean item response latencies and the simulated response latencies was .04 (*ns*), and in Experiment 2B, this correlation was essentially zero (.008). What these observations suggest is that although the model can be configured in such a way that the inhibitory effect of neighborhood frequency can be eliminated (as in our simulations), or perhaps even reversed, the way that this is realized in the model (at least in our simulations) bears little resemblance to the way that participants are actually responding in the experiment.

Discussion

In these experiments, the nonwords were matched to the words on neighborhood size. As was the case in Experiment 1C, in these situations, it should not be possible for participants to use global lexical activation as a reliable cue for responding. According to the multiple read-out model, then, responses should be made following lexical selection (i.e., when the M criterion is exceeded), and the model thus predicts that a large inhibitory neighborhood frequency effect should be observed. This prediction was not borne out in either of these experiments. Instead, as was the case in Experiment 1C, the neighborhood frequency effect was facilitatory, not inhibitory. Simulations using the multiple read-out model indicated that if only the M criterion was used for responding, then the model would predict an inhibitory neighborhood frequency effect in both experiments. This predicted inhibitory effect of higher frequency neighbors could be eliminated when the Σ and T criteria were incorporated into the simulations, but a clear facilitatory effect of higher frequency neighbors could not be simulated unless the simulated nonword error rate was substantially larger than that observed in the experiments.

General Discussion

The primary motivation for this research was to examine the multiple read-out model's account of neighborhood size and neighborhood frequency effects in lexical decision tasks as a function of nonword orthographic neighborhood size. As noted, in Experiments 1A–1D, the nonword context was more extensively manipulated than in Grainger and Jacobs's (1996) experiments. Specifically, Grainger and Jacobs used two nonword contexts (the nonwords had either small neighborhoods or large neighborhoods), whereas in the present experiments, there were four nonword contexts. That is, in Experiment 1A, the nonwords had no neighbors, in Experiment 1B, the nonwords had small neighborhoods, in Experiment 1C, the nonwords were matched to the words on neighborhood size, and in Experiment 1D, the nonwords had large neighborhoods. This more extensive manipulation of the nonword context allowed for a more comprehensive test of the model's predictions regarding the effects of orthographic neighbors in the lexical decision task. In addition, the word frequency manipulation allowed us to gauge the extent to which lexical selection (i.e., the M criterion in the multiple read-out model) was involved during responding in each experiment, which, as previously noted, should be related to both the magnitude of the facilitatory neighborhood size effect and the magnitude and direction of the neighborhood frequency effect.

Considered together, the multiple read-out model's predictions were not well supported by the data. In fact, the only case in which the model's predictions were unequivocally supported was in Experiment 1A. In that experiment, the model predicted a facilitatory neighborhood size effect and a facilitatory neighborhood frequency effect, both of which were observed. In Experiment 1B, the model predicted a facilitatory neighborhood size effect and, for the small neighborhood words, an inhibitory neighborhood frequency effect. Although the neighborhood size effect was facilitatory, the effect of neighborhood frequency was also facilitatory for both the small neighborhood words and the large neighborhood words (both effects interacted with word frequency, as they were observed for only the low-frequency words). In Experiment 1C, the model also predicted no effect of neighborhood size and an inhibitory neighborhood frequency effect. Instead, both effects were facilitatory for the low-frequency words, whereas there were no effects of neighborhood size or of neighborhood frequency for the high-frequency words. In Experiment 1D, the model predicted no neighborhood size effect and an inhibitory neighborhood frequency effect. In fact, for the low-frequency words, there was no effect of neighborhood size; however, there was some evidence of an inhibitory neighborhood size effect for the high-frequency words. Moreover, the effect of neighborhood frequency was facilitatory, not inhibitory.

The model's predictions regarding the neighborhood frequency effect were tested further in Experiments 2A and 2B. In these experiments, neighborhood size was controlled and neighborhood frequency was manipulated, creating a situation in which the model predicts that an inhibitory neighborhood frequency effect will be observed. However, the neighborhood frequency effect was again facilitatory, both for words with small neighborhoods and for words with large neighborhoods.

Of all these results, two are particularly troublesome for the multiple read-out model. The first is the facilitatory neighborhood size effect observed in Experiment 1C, in which the words and the nonwords were matched on neighborhood size. According to the model, if the words and the nonwords cannot be distinguished from one another on the basis of lexical activation (i.e., via use of the Σ criterion), then a neighborhood size effect should not be observed. That is, the model assumes that a facilitatory neighborhood size effect will occur only when participants base their responses on global lexical activity (i.e., when they use the Σ criterion for responding). Our simulations indicated that in this experiment, the degree of lexical activation was not useful for distinguishing the words from the nonwords. This prohibits the extensive use of the Σ criterion necessary to produce a facilitatory neighborhood size effect. Thus, the facilitatory neighborhood size effect witnessed in this experiment casts serious doubt on the model's assumption that neighborhood size effects are strictly a global activation phenomenon (i.e., due to the Σ criterion, as opposed to the M criterion).

The second problematic finding for the model is the lack of any evidence for an inhibitory neighborhood frequency effect. Because the model assumes that inhibitory neighborhood frequency effects are due to lexical selection processes (i.e., intralevel inhibition between word units), inhibitory neighborhood frequency effects should have been observed whenever the words could not be reliably distinguished from the nonwords on the basis of lexical activation (i.e., when the nonwords had large neighborhoods or when the words and the nonwords were matched on neighborhood size). In Experiment 1D, the nonwords had large neighborhoods, and in Experiments 1C and 2, the words and the nonwords were matched on neighborhood size; yet in all of these experiments, facilitatory neighborhood frequency effects were observed. Simulations with the multiple read-out model confirmed that the degree of lexical activation was not useful for distinguishing the words from the nonwords in these experiments, and, therefore, inhibitory (or at least null) effects of neighborhood frequency should have been observed. These results cast serious doubt on the model's assumption that the inhibitory neighborhood frequency effect is a lexical selection effect and hence is based on mechanisms intrinsic to word recognition.

Of course, this conclusion would seem to be incompatible with the numerous reports of inhibitory neighborhood frequency effects in lexical decision studies (e.g., Carreiras et al., 1997; Grainger, 1990; Grainger & Jacobs, 1996; Grainger et al., 1992; Grainger & Segui, 1990; Huntsman & Lima, 1996; Perea & Pollatsek, 1998). As Andrews (1997) noted, however, inhibitory neighborhood frequency effects have not been consistently observed when English stimuli are used, and in some cases, those effects are facilitatory (as was the case in the present experiments; see Forster & Shen, 1996; Sears et al., 1995). It is also worth noting that facilitatory neighborhood frequency effects for English words have been observed in perceptual identification tasks (Sears, Lupker, & Hino, 1999) as well. This fact is important because according to the multiple read-out model, in perceptual identification tasks, the Σ criterion cannot be used for responding (because a word must be uniquely identified before an accurate response can be made), and thus, responses should be based solely on the M criterion. Consequently, the inhibitory effects of higher frequency neighbors should be very clear in this task, as they have been in studies that have used French stimuli (e.g., Grainger & Jacobs, 1996) and Spanish stimuli (Carreiras et al., 1997). Such was not the case, however, in the Sears, Lupker, and Hino (1999) experiments. Together, these observations suggest that there may be important language differences in the role that inhibition plays in orthographic processing, with the inhibitory process being more dominant in some languages (French and Spanish) than in others (English).

Recall that Andrews (1997) proposed that the reason facilitatory effects of neighborhood size are commonly observed in English, but not in French or Spanish, is because the word body (i.e., the orthographic rime) may play a special role when reading English words. Consistent with this proposal, Ziegler and Perry (1998) reported that responses to English words with many body neighbors were faster than responses to words with few body neighbors. Facilitatory effects of higher frequency neighbors for English words could have a similar origin. That is, for English words, the higher frequency neighbors of a word will often be body neighbors, and those body neighbors, by virtue of their higher frequency, may have a strong influence on the lexical selection process. In any case, it is becoming apparent that there are important differences among languages in the role that inhibition plays during orthographic processing (e.g., Van Heuven, Dijkstra, & Grainger, 1998).

A consideration of the neighborhood statistics of the English language may be instructive in this regard. Andrews (1997) reported that in a sample of 1,895 four-letter words, 80.3% had one or more higher frequency neighbors. For five- and six-letter words, which have much smaller neighborhoods than four-letter words, the corresponding percentages were 52.0% and 30.2%, respectively. (Of the 519 three-letter words listed in the Kucera and Francis [1967] norms, 87.8% have at least one higher frequency neighbor.) These statistics indicate that the majority of English words three to five letters in length have higher frequency neighbors. In our view, a lexical processor that delays the processing of the majority of words and facilitates the processing of the minority is, at best, counterintuitive.

The Role of Orthographic Neighbors in Other Models of Word Recognition

The role of orthographic neighbors in activation-based models such as the multiple read-out model and the interactive-activation model has been discussed at length. But what of other models of word recognition and their predictions regarding the effects of orthographic neighbors? As Andrews (1997) has pointed out, models that incorporate a serial-search mechanism (e.g., Forster, 1976; Paap et al., 1982; but see Forster, 1989) have difficulties accounting for facilitatory neighborhood size effects and facilitatory neighborhood frequency effects. In these models, the presentation of a word activates a candidate set of word entries (i.e., words that are orthographically similar to the presented word), and higher frequency words in the candidate set are checked before lower frequency words, with the search continuing until a correct match is found (at which point word identification is achieved). Because the search is frequency ordered, responses to words with higher frequency neighbors (and typically to words with large neighborhoods, because many low-frequency words with large neighborhoods have higher frequency neighbors) will be slower than responses to words without higher frequency neighbors (and to words with small neighborhoods). Thus, these models predict an inhibitory neighborhood frequency effect, and typically an inhibitory neighborhood size effect, for low-frequency words, as low-frequency words are more likely to have higher frequency neighbors. Clearly, these models are unable to account for the data of the present study, as large neighborhoods and higher frequency neighbors facilitated responses to low-frequency words. (Forster's [1989] recent version of the serial-search model no longer predicts inhibitory neighborhood size or inhibitory neighborhood frequency effects for words. In this revised model, neither neighborhood size nor neighborhood frequency are predicted to have any effect on word identification latencies.)

Recently, Sears, Hino, and Lupker (1999) have examined the predictions of parallel distributed processing models (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) with regard to orthographic neighborhood effects. Unlike the multiple read-out model and the interactive-activation model, in these models, there are no lexical units that represent single words. Instead, lexical representations are embodied in the pattern of activation across an interconnected network of units. In a series of statistical analyses of the orthographic, phonological, and cross-entropy error scores of the four- and five-letter monosyllabic words in these models' corpi, it was found that for low-frequency words, words with large neighborhoods and words with higher frequency neighbors had, on average, lower error scores than words with small neighborhoods and words with no higher frequency neighbors. Because lower error scores correspond to faster lexical decision and pronunciation latencies in these models, the models, therefore, predict the finding that large neighborhoods and higher frequency neighbors facilitate responses to low-frequency words.

Conclusions

The purpose of this research was to examine the multiple read-out model's account of neighborhood size and neighborhood frequency effects in lexical decision tasks as a function of nonword orthographic neighborhood size. In contrast to the predictions of the model, the present results show that large neighborhoods and higher frequency neighbors facilitate responses to low-frequency words in a wide variety of nonword contexts. Thus, the multiple read-out model, as currently instantiated, is not a viable account of orthographic neighborhood effects, at least for the processing of English words.

References

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.

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

Bozon, F., & Carbonnel, S. (1996). Influence des voisins orthographiques sur l'identification de mots et de pseudomots [Orthographic neighborhood effect in word and pseudoword recognition]. *Revue de Neuropsychologie, 6,* 219–237.

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

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28,* 143–153.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Cohen, J. (1976). Random means random. *Journal of Verbal Learning and Verbal Behavior, 15,* 261–262.

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

Content, A. (1991). The effect of spelling-to-sound regularity on naming in French. *Psychological Research, 53,* 3–12.

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

Forster, K. I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and processing* (pp. 75–107). Cambridge, MA: MIT Press.

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

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228–244.

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

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics, 45,* 189–195.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics, 51,* 49–56.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition. *Perception & Psychophysics, 47,* 191–198.

Huntsman, L. A., & Lima, S. D. (1996). Orthographic neighborhood structure and lexical access. *Journal of Psycholinguistic Research, 25,* 417–429.

Johnson, N. F., & Pugh, K. R. (1994). A cohort model of visual word recognition. *Cognitive Psychology, 26,* 240–346.

Keppel, G. (1976). Words as random variables. *Journal of Verbal Learning and Verbal Behavior, 15,* 263–265.

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

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General, 122,* 316–330.

Mathey, S., & Zagar, D. (1996). Rôle du voisinage orthographique lors de la reconnaissance visuelle des mots de 4, 6, et 8 lettres [The role of orthographic neighborhood in the visual recognition of 4-, 6-, and 8-letter words]. *Revue de Neuropsychologie, 6,* 205–217.

Mathey, S., & Zagar, D. (2000). The role of neighborhood distribution in visual word recognition: Words with single and twin neighbors. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 184–205.

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

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

Peereman, R., & Content, A. (1995). Neighborhood size effect in naming words and pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 409–421.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 767–779.

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

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

Pynte, J. (2000). Words likely to activate many lexical candidates are granted an advantage: Evidence from within-word eye movements. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 47–64). Oxford, United Kingdom: Elsevier.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416–429.

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.

Sears, C. R., Hino, Y., & Lupker, S. J. (1999). Orthographic neighborhood effects in parallel distributed processing models. *Canadian Journal of Experimental Psychology, 53,* 220–229.

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

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

Snodgrass, J. G., & Minzer, M. (1993). Neighborhood effects in visual word recognition: Facilitatory or inhibitory? *Memory & Cognition, 21,* 247–266.

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

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.

Wike, E. L., & Church, J. D. (1976). Comments on Clark's "the language as fixed-effect fallacy." *Journal of Verbal Learning and Verbal Behavior, 15,* 249–255.

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

Appendix

Stimuli Used in the Experiments

Items Used in Experiments 1A–1D

High-Frequency Words With a Small Neighborhood and No Higher Frequency Neighbors

ABLE, ARMY, BLUE, CLUB, DATA, DESK, GIRL, HUGE, STEP, TRUE, TYPE, UNIT, VIEW, BIRTH, BLOOD, CHECK, CHEST, COAST, DOZEN, DREAM, DRINK, LOOSE, METAL, PHONE, SPOKE, and STYLE

High-Frequency Words With a Small Neighborhood and Higher Frequency Neighbors

DOWN, EASY, FAIR, FIRM, JOIN, RISK, SIZE, SOFT, SPOT, TEXT, TREE, VOTE, WALK, ALONE, BEGIN, CLEAN, DEPTH, HEART, IDEAL, MOUTH, PEACE, THICK, WOMEN, WORTH, WRITE, and YOUTH

High-Frequency Words With a Large Neighborhood and No Higher Frequency Neighbors

BOAT, BORN, CALL, FLAT, FLOW, MAIN, PAGE, PICK, RISE, ROCK, ROLE, STAY, TEAM, BREAK, CARRY, CLASS, PARTY, REACH, RIVER, SCALE, SHARE, SHORT, SPITE, SWEET, TRAIN, and WATCH

High-Frequency Words With a Large Neighborhood and Higher Frequency Neighbors

CLAY, COOL, DATE, FOOT, LAND, LATE, NOSE, PASS, RACE, REST, SEND, WARM, WIDE, EIGHT, FIGHT, GRASS, HORSE, LOWER, ROUND, SCORE, SHAPE, SHORE, SIGHT, SOUND, STAGE, and STORE

Low-Frequency Words With a Small Neighborhood and No Higher Frequency Neighbors

CRIB, DEBT, DIRT, DUMB, FUSE, GASP, GLAD, HURT, INCH, LION, MONK, STUD, TUBE, BLAME, BLAST, BOOST, BRICK, CRAWL, GLAZE, GLOOM, HARSH, PLEAD, SAUCE, SLAVE, SPRAY, and STEEL

Low-Frequency Words With a Small Neighborhood and Higher Frequency Neighbors

CALF, CLUE, FOAM, FUEL, GOWN, HORN, KNEE, KNOT, PITY, SHUT, STEM, TWIN, VERB, BLOND, BLOWN, LOYAL, REACT, SKILL, SHOOT, SPADE, SPORT, STEAK, STUFF, TREAT, WEAVE, and YIELD

Low-Frequency Words With a Large Neighborhood and No Higher Frequency Neighbors

BOWL, CUTS, DUKE, JUMP, LOAN, PATH, PLOT, PUSH, RAFT, RIBS, SKIN, SLAB, SPAN, BAKER, BORED, GRACE, JOLLY, LUNCH, METER, PITCH, PORCH, SCOUT, SHINE, SILLY, TIRED, and WIPED

Low-Frequency Words With a Large Neighborhood and Higher Frequency Neighbors

BENT, BOOM, CAPE, CAST, CURE, CORD, GATE, GAZE, HALT, MALL, NEST, PACE, RICE, BLANK, BOUND, FIRED, GRADE, PEACH, PLATE, POKER, SHADE, SLACK, SPICE, SPIKE, SPILL, and TRACE

Items Used in Experiments 2A and 2B

Small Neighborhood Words With No Higher Frequency Neighbors

CRIB, DEBT, DIRT, DRUG, DUMB, EXIT, FUSE, GASP, GLAD, HURT, INCH, LION, MONK, STUD, TUBE, BLAME, BLAST, BOOST, BRICK, CRASH, CRAWL, GLAZE, GLOOM, HARSH, PLEAD, SAUCE, SLAVE, SPRAY, STAMP, and STEEL

Small Neighborhood Words With One Higher Frequency Neighbor

BAIT, CALM, COMB, DISH, FISH, FOAM, GENE, HORN, KNEE, KNOT, MOTH, PITY, SHUT, TWIN, VERB, BLOND, BLOWN, COUNT, FEAST, LOYAL, REACT, SHOOT, SKILL, SPADE, SPORT, STEAK, STUFF, TREAT, WEAVE, and YIELD

Large Neighborhood Words With No Higher Frequency Neighbors

BOWL, CUTS, DUKE, JUMP, LOAN, PATH, PEAS, PLOT, PUSH, RAFT, RIBS, SKIN, SLAB, SPAN, SUMS, BAKER, BORED, EAGER, FREED, GRACE, JOLLY, LUNCH, METER, PITCH, PORCH, SCOUT, SHINE, SILLY, TIRED, and WIPED

Large Neighborhood Words With One Higher Frequency Neighbor

CASH, CODE, CORN, KISS, LASH, LEAF, LINK, LOOP, MAIL, MAPS, MINK, ROLL, TACT, TART, WASH, BLANK, FIRED, GRADE, GROWN, HONEY, LAYER, PAINT, POKER, PRIME, SLACK, SMELL, SPICE, SPILL, STARS, and TRACE

Received October 13, 1999
Revision received July 11, 2001
Accepted September 11, 2001 ■