UNIVERSITY OF CALGARY

The Effects of Nonword Orthographic Neighborhood Size in Lexical Decision

by

Paul D. Siakaluk

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

July 19, 1999

0-612-48042-9

Abstract

In eight lexical decision experiments, the effect of large neighborhoods (neighborhood size) and the effect of higher frequency neighbors (neighborhood frequency) were examined as a function of nonword orthographic neighborhood size. According to the multiple read-out model (Grainger & Jacobs, 1996), the neighborhood size effect and the neighborhood frequency effect will vary systematically as a function of the nonword neighborhood size. In the present experiments the nonword context was more extensively manipulated than in previous studies to further test the model's predictions and explanations for the neighborhood size effect and the neighborhood frequency effect. Overall, the results of this study are very problematic for the model, but by changing one of the model's critical assumptions many of the results could be accommodated.

Acknowledgements

In the past two years I have learned a great deal about research in cognitive psychology from my supervisor Chris Sears, and I wish to thank him for his advice and support. I also wish to thank Penny Pexman and Chip Scialfa for their time and help with this research project, and to Hotze Rullmann for serving on my examination committee.

I am grateful for the friendship and support of Geoff Ho, Jayne Hutchinson, Alisa McArthur, Lenora Brown, Connie Legare, Simon Chi, and Yan Wong. I am also grateful for the support I received from the other faculty and graduate students in the Department of Psychology. Finally, I wish to thank my family, especially my wife, children, and parents, for their patient support during the past two years.

TABLE OF CONTENTS

LIST OF TABLES

The Effects of Nonword Orthographic Neighborhood Size in Lexical Decision

Introduction

In the past decade a considerable amount of research has sought to determine 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). The most commonly used orthographic similarity metric is Coltheart's N (Coltheart, Davelaar, Jonasson, & Besner, 1977), which is defined as the number of different words that can be created by changing one letter of a word while maintaining letter positions. For example, the words FREE, THEE, TRUE, and TREK are all orthographic neighbors of the word TREE.

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. The neighborhood sizes of words varies considerably, as some words have a large neighborhood (e.g., MALE, with twenty-two neighbors), others have a small neighborhood (e.g., GIRL, with two neighbors), and others have no neighbors at all (e.g., IDOL). The second characteristic of interest is the existence of higher frequency neighbors in the word's orthographic neighborhood, usually referred to as neighborhood frequency. The word PLOT, for example, has a Kucera and Francis (1967) normative frequency per million words of 37, and the normative frequency of its highest frequency neighbor (PLOW) is 12. Thus, PLOT has no higher frequency neighbors. In contrast, the word LIME has many higher frequency neighbors. Specifically, LIME has a normative frequency of 13, and the normative frequencies of its highest frequency neighbors (TIME, LIKE, LIFE, LINE, and LIVE) are 1599, 1290, 715, 298, and 177, respectively.

For many models of visual word recognition, the number of neighbors and the existence of higher frequency neighbors will 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 will then play a role in the lexical selection (i.e., word identification) process. The precise role that orthographic neighbors play in word identification has been primarily investigated using the lexical decision task, and several contradictory findings have emerged from this body of empirical research.[1]

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 are responded to more rapidly than words with small neighborhoods (a facilitatory neighborhood size effect). Importantly, this facilitatory neighborhood size effect is usually only observed for 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 to high-frequency words. Facilitatory neighborhood size effects for low-frequency words in lexical decision tasks have also been reported by Forster and Shen

(1996), and Sears, Hino, and Lupker (1995), with the latter 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 (Carreiras, Perea, & Grainger, 1997; Grainger & Jacobs, 1996; Johnson & Pugh, 1994). Grainger and Jacobs 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. intermixed small and large neighborhood nonwords in the lexical decision task, and reported no effect of neighborhood size. Finally, Johnson and Pugh 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. (When their word and nonword stimuli were not blocked by neighborhood size, however, they reported a trend towards facilitation.)

Taken together, Andrews (1997) argues that 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 consistently 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), or when the words and nonwords are blocked by neighborhood size (Johnson & Pugh, 1994).

Andrews (1997) offered a language-specific explanation for why facilitatory effects of neighborhood size are observed in English, but not in French or Spanish. English has an inconsistent relationship between orthography and phonology, and the word body (i.e., the orthographic rime) may play a special role when reading English words, because vowels are more inconsistently pronounced than consonants, and consonants that follow a vowel better predict its correct pronunciation than consonants that precede it (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). Thus, the word body may be an important functional unit in the English orthographic lexicon. Word bodies 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 the 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). (Note that a word was not restricted to be of the same length as the target word to be considered a body neighbor; i.e., the word FEED was considered to be a body neighbor of the target word BLEED).

Ziegler and Perry's results suggest that the neighborhood size effect for English words may be due to the fact that 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. 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. It is also worth noting that facilitatory neighborhood size effects have been observed with French words (Grainger & Jacobs, 1996, Experiment 1C), and so a completely language-specific explanation of neighborhood size effects would seem to be inadequate.

Effects of Higher Frequency Neighbors

A number of investigators have examined the effects of a word's higher frequency neighbors on its identification latencies (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 seem 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), neighborhood frequency effects are usually inhibitory in the lexical decision task. Andrews (1997) also noted 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). Overall, 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' (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 (1996) have proposed an activation-based model which can apparently accommodate both facilitatory neighborhood size effects and inhibitory (as well as facilitatory) neighborhood frequency effects in lexical decision tasks.[2] Grainger and Jacobs' "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. Intra-level 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 is intended to explain inhibitory neighborhood frequency effects.

The multiple read-out model is unique from the interactive-activation model in that it also incorporates three variable decision criteria which 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. 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.[3]

The setting of the M criterion is fixed in the model, whereas the setting of the Σ and T criteria can be strategically adjusted, based on 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 will determine whether positive responses are based more on lexical selection or on 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 1) the facilitatory neighborhood size effect (and the facilitatory neighborhood frequency effect) in lexical decision do not actually arise during the lexical selection process (i.e., the M criterion being exceeded), but instead occur when participants use the Σ criterion for responding, and 2) the inhibitory neighborhood frequency effect is a true lexical selection effect, resulting from the intra-level competitive processes which occur prior to the M criterion being exceeded.

To test these assumptions, Grainger and Jacobs (1996) conducted two lexical decision experiments in which the neighborhood sizes of the nonwords was varied. In these experiments, for the word stimuli neighborhood size and neighborhood frequency were manipulated by using 1) words with few neighbors, none of higher frequency, 2) words with few neighbors, one of higher frequency, 3) words with many neighbors, of which one was of higher frequency, and 4) 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 (effect of one higher frequency neighbor), conditions 2 and 3 (effect of neighborhood size), and conditions 3 and 4 (effect of number of higher frequency neighbors).

In Grainger and Jacob's (1996) Experiment 1C, the nonwords all had small neighborhoods. According to the multiple read-out model, in this situation the Σ criterion should generally be set below 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). If the participants use the Σ criterion to respond, then words with large neighborhoods will be easier to distinguish from the nonwords than the words with small neighborhoods, because words with large neighborhoods should generally produce more lexical activity than words with small neighborhoods. Consequently, a facilitatory neighborhood size effect should be observed. 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. This was the case in Grainger and Jacob's Experiment 1C, and simulations with the model indicated it was successful in simulating these effects.

In Grainger and Jacobs' (1996) Experiment 1B, the nonwords all had large neighborhoods, and in this situation, according to the model, the Σ criterion should be set higher than the M criterion 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 only occur 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 the nonwords all 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 the model was successful in simulating this outcome as well.

Overall, Grainger and Jacobs (1996) argued that the multiple read-out model could account for many of the conflicting findings in the literature regarding the effects of neighborhood size and neighborhood frequency. In the present study, the model's ability to account for such effects when the nonword context is more extensively manipulated was examined.

Experiment 1

As in Grainger and Jacobs' (1996) study, in Experiments 1A-1D the neighborhood sizes of the nonwords used in lexical decision tasks was manipulated to test the multiple read-out model's account of neighborhood size and neighborhood frequency effects in this task. The present experiments differed from those of Grainger and Jacobs, however, in several important ways. First, in the present experiments word frequency (high, low), neighborhood size (small, large), and neighborhood frequency (no higher frequency neighbors, higher frequency neighbors) were factorially manipulated.

(The same set of words was used in each of the experiments). This design should better clarify how neighborhood size and neighborhood frequency affect the processing of both high-frequency and low-frequency words, and allow for tests of any interactions between these factors.

Second, and more importantly, in Experiments 1A-1D the nonword context was more extensively manipulated than in Grainger and Jacobs' (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; namely, the nonwords had no neighbors, small neighborhoods, large neighborhoods, or were matched to the words on neighborhood size. This more extensive manipulation of the nonword context allows for a more comprehensive test of the model's assumptions and its 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 so 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 thus set their Σ criterion below the M criterion for most (if not all) of the 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 others.

In Experiment 1B the nonwords had small neighborhoods (e.g., GRUN, with two neighbors). According to the model, in this experiment, because the words with large neighborhoods will generate more global lexical activity than the small neighborhood words and nonwords, participants will generally set their Σ criterion below the M criterion, in which case a facilitatory neighborhood size effect should be observed. Because the words with small neighborhoods and the nonwords with small neighborhoods cannot be reliably distinguished from one another via the Σ criterion, however, the model further assumes that some responses will be based upon the M criterion being exceeded, in which case an inhibitory neighborhood frequency effect should be observed. Note that this experiment is similar to Grainger and Jacobs' (1996) Experiment 1C, and thus serves as an attempted replication of their findings with a different and more extensive set of stimuli.

In Experiment 1C the nonwords had large neighborhoods (e.g., DAST, with 13 neighbors), as was the case in Grainger and Jacobs' (1996) Experiment 1B. Recall that, according to the model, because the words with large neighborhoods will not generate more lexical activity than the nonwords with large neighborhoods, 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 higher than the M criterion, and because most 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 relative to its use in Experiment 1B. The model also predicts that the neighborhood size effect should be slightly facilitatory or non-existent, because the Σ criterion should be only used occasionally for responding.

In Experiment 1D, 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 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 greatest in this experiment, because more responses should be based on the M criterion in this experiment than in any of the others. The predictions of the model for each of these experiments are summarized in Table 1.

General Method

Participants. One-hundred and sixty undergraduate students from the University of Calgary participated in the experiments: forty 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

Table 1

Qualitative Predictions for Neighborhood Size Effects (N), Neighborhood Frequency Effects (NBF), and Word Frequency Effects (WF) in Experiments 1A-1D

		Effect		
Experiment	Criterion Setting	N	NBF	WF
1A (ZNB nonwords)	$\Sigma < M$	Facilitatory (X)	Facilitatory (X)	Small
1B (SNB nonwords)	$\Sigma < M$	Facilitatory (X)	Inhibitory (x)	Medium
1C (LNB nonwords)	$\Sigma > M$	Facilitatory (x)	Inhibitory (X)	Large
1D (MNB nonwords)	$\Sigma > M$	Null	Inhibitory (X)	Large

Note. ZNB = zero neighbors; SNB = small neighborhood; LNB = large neighborhood; MNB = nonwords matched to words on neighborhood size. < refers to lower setting of the Σ criterion relative to the M criterion. > refers to higher setting of the Σ criterion relative to the M criterion. X = large effect; x = small effect.

experiment.

Stimuli. The complete set of experimental words used in Experiments 1A, 1B, 1C, and 1D is presented in Appendix A, and the descriptive statistics for these stimuli are listed in Table 2.

Half the words 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 and Francis (1967) normative frequency per million words of 105.6 (range of 52 to 231), and the remainder of the words were low-frequency words with a mean normative frequency per million words of 21.3 (range of 1 to 48).

The second factor manipulated was neighborhood size. Half of the words had small neighborhoods (i.e., at least one neighbor and no more than 5 neighbors); these had a mean neighborhood size of 3.3. The other half of the words had large neighborhoods (i.e., at least 6 neighbors; range of 6 to 18); these had a mean neighborhood size of 9.8. 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 that was of higher frequency than themselves, and the other half of the words 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 292.2. For the low-frequency words with higher frequency neighbors, the mean Kucera and Francis

Table 2

Mean Word Frequency and Neighborhood Size (N) for the Word Stimuli Used in Experiments 1A-1D

	Low-frequency words		High-frequency words	
Neighborhood frequency	Small N	Large N	Small N	Large N
No HFN				
Frequency	19.1	23.8	114.4	110.1
N	3.3	8.7	3.1	9.4
NBF	8.0	13.6	27.7	55.0
HFN				
Frequency	19.4	20.3	105.7	104.2
N	3.4	10.1	3.2	10.7
NBF	273.9	310.6	280.7	299.6

Note. HFN = higher frequency neighbors. NBF refers to the average frequency of the highest frequency neighbor.

normative frequency of the highest frequency neighbor of each word was 293.5. 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.

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 of 1 to 5 neighbors, with a mean neighborhood size of 3.4), and in Experiment 1C the nonwords had large neighborhoods (range of 6 to 21 neighbors, with a mean neighborhood size of 9.8). In Experiment 1D, the nonwords were matched closely to the words on neighborhood size. More specifically, 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. The mean neighborhood sizes of the large neighborhood words and the large neighborhood nonwords were identical (9.8). (The overall mean neighborhood size of the nonwords used in Experiment 1D was 6.5.)

Apparatus and procedure. Stimuli were presented on a color VGA 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 1s 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 1 s.

Each participant completed 16 practice trials prior to the collection of data. The practice stimuli consisted of eight words (four of low-frequency, 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 experiment proper, 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 (percent error), 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, low) x 2 (Neighborhood Size: small, large) x 2 (Neighborhood Frequency: no higher frequency neighbors, 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, 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, low) x 2 (Neighborhood Size: small, large) x 2

(Neighborhood Frequency: no higher frequency neighbors, higher frequency neighbors) repeated-measures factorial analysis of variance (ANOVA). Both subject (F_s) and item (F_i) analyses were carried out.[4]

Experiment 1A: Lexical Decision with Nonwords with No Neighbors

Method

Participants, stimuli, design, and procedure. The participants, word stimuli, design, and procedure are described in the General Method. The nonword stimuli consisted of 104 four-letter and 104 five-letter, orthographically-legal and pronounceable letter strings. The nonwords had no orthographic neighbors.

Results

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. A total of 4 observations (0.04% of the data) were removed by this procedure. The mean response latencies of correct responses and the mean error rates are shown in Table 3. The mean response latency and percent error for the nonwords were 540 ms and 2.5%, respectively.

Response latencies. There was a significant main effect of word frequency, $F_s(1, 39) = 77.55$, $p < .001$, $MSE = 622.37$; $F_i(1, 200) = 34.74$, $p < .001$, $MSE = 934.02$, as responses to high-frequency words were an average of 24.5 ms faster than responses to low-frequency words. The main effect of neighborhood size was significant, $F_s(1, 39) = 23.07$, $p < .001$ $MSE = 218.17$; $F_i(1, 200) = 3.96$, $p < .05$, $MSE = 934.02$, as was the main effect of neighborhood frequency, $F_s(1, 39) = 15.48$, $p < .001$, $MSE = 307.73$; $F_i(1, 200) = 3.94$, $p < .05$, $MSE = 934.02$. Responses to words with large neighborhoods were an average of 8 ms faster than responses to words with small neighborhoods, and responses

Table 3

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 1A

	Neighborhood size	
Neighborhood frequency	Small	Large
	Low-frequency words	
No HFN	514 (4.4)	506 (2.5)
HFN	504 (2.7)	493 (1.3)
	High-frequency words	
No HFN	486 (1.7)	477 (0.7)
HFN	480 (1.6)	476 (1.4)

Note. HFN = higher frequency neighbors. Error rates in parenthesis ().

to words with higher frequency neighbors were an average of 7.5 ms faster than responses to words without higher frequency neighbors.

The interaction between word frequency and neighborhood frequency was significant, $F_s(1, 39) = 5.03$, $p < .05$, MSE = 258.16; $F_i < 1$. For low-frequency words, the neighborhood frequency effect was facilitatory – responses to words with higher frequency neighbors were an average of 11.5 ms faster than responses to words with no higher frequency neighbors, $F_s(1, 39) = 20.61$, $p < .001$, MSE = 267.85; $F_i(1, 100) = 2.97$, $p = .088$, MSE = 1,372.66. For high-frequency words, responses to words with higher frequency neighbors were an average of 3.5 ms faster than responses to words with no higher frequency neighbors ($F_s(1, 39) = 1.82$, $p = .185$, MSE = 298.04; $F_i < 1$). No other interactions were significant (all F's < 1).

Error rates. The main effect of word frequency was significant, $F_s(1, 39) = 27.82$, $p < .001$, MSE = 5.59; $F_i(1, 200) = 12.29$, $p < .005$, MSE = 7.94, as was the main effect of neighborhood size, $F_s(1, 39) = 10.06$, $p < .005$, MSE = 9.72; $F_i(1, 200) = 9.09$, $p < .005$, MSE = 7.94. Participants made fewer errors to high-frequency words than to low-frequency words (1.3% vs. 2.7%) and fewer errors to words with large neighborhoods than to words with small neighborhoods (1.4% vs. 2.6%). The main effect of neighborhood frequency was marginally significant, $F_s(1, 39) = 3.70$, $p = .062$, MSE = 7.19; $F_i(1, 200) = 2.76$, $p = .098$, MSE = 7.94. Participants generally made fewer errors to words with higher frequency neighbors than to words with no higher frequency neighbors (1.7% vs. 2.3%).

As was the case in the response latency analysis, the word frequency by neighborhood frequency interaction was significant, $\underline{F}_s(1, 39) = 9.75$, $\underline{p} < .005$, $\underline{MSE} = 6.14$; $\underline{F}_i(1, 200) = 5.18$, $\underline{p} < .05$, $\underline{MSE} = 7.94$. For the low-frequency words, words with higher frequency neighbors were responded to more accurately than words with no higher frequency neighbors (2.0% vs. 3.4%), $\underline{F}_s(1, 39) = 13.00$, $\underline{p} < .005$, $\underline{MSE} = 6.40$; $\underline{F}_i(1, 100) = 5.64$, $\underline{p} < .05$, $\underline{MSE} = 10.92$). For the high-frequency words, there was no effect of neighborhood frequency (both $\underline{F}$'s < 1). The word frequency by neighborhood size interaction was marginally significant, $\underline{F}_s(1, 39) = 3.23$, $\underline{p} = .080$, $\underline{MSE} = 6.91$; $\underline{F}_i(1, 200) = 2.00$, $\underline{p} = .159$, $\underline{MSE} = 7.94$. For the low-frequency words, words with large neighborhoods were responded to more accurately than words with small neighborhoods (1.9% vs. 3.5%), $\underline{F}_s(1, 39) = 9.01$, $\underline{p} < .01$, $\underline{MSE} = 11.86$; $\underline{F}_i(1, 100) = 7.13$, $\underline{p} < .01$, $\underline{MSE} = 10.92$. For the high-frequency words, although words with large neighborhoods were responded to more accurately than words with small neighborhoods (1.0% vs. 1.6%), this effect was not statistically significant, $\underline{F}_s(1, 39) = 2.78$, $\underline{p} = .103$, $\underline{MSE} = 4.78$; $\underline{F}_i(1, 100) = 2.05$, $\underline{p} = .156$, $\underline{MSE} = 4.96$). No other interactions were significant (all $\underline{F}$'s < 1).

Discussion

The nonwords used in Experiment 1A had no neighbors, and thus they would generate 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 will 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 note. 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. 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.

Experiment 1B: Lexical Decision with Small Neighborhood Nonwords

Method

Participants, stimuli, design, and procedure. The participants, word stimuli, design, and procedure are described in the General Method. The nonword stimuli consisted of 104 four-letter and 104 five-letter, orthographically-legal and pronounceable letter strings. The nonwords all had small neighborhoods (range of 1 to 5 neighbors, with a mean neighborhood size of 3.4).

Results

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. A total of 5 observations (0.06% of the data) were removed by this procedure. The mean response latencies of correct responses and the mean error rates are shown in Table 4. The mean response latency and percent error for the nonwords were 607 ms and 4.5%, respectively.

Response latencies. The main effect of word frequency was significant, $F_s(1, 39) = 161.46$, $p < .001$, $MSE = 777.71$; $F_i(1, 200) = 59.52$, $p < .001$, $MSE = 1,550.76$. Responses to high-frequency words were an average of 39.5 ms faster than responses to low-frequency words. The main effect of neighborhood size and the main effect of neighborhood frequency were also significant, $F_s(1, 39) = 9.12$, $p < .005$, $MSE = 390.37$; $F_i(1, 200) = 1.87$, $p = .173$, $MSE = 1,550.76$, and $F_s(1, 39) = 4.71$, $p < .05$, $MSE = 585.09$; $F_i(1, 200) = 1.29$, $p = .257$, $MSE = 1,550.76$, respectively. Responses to words with large neighborhoods were an average of 6.5 ms faster than responses to words with small neighborhoods, and responses to words with higher frequency neighbors were an average of 6 ms faster than responses to words with no higher frequency neighbors.

The interaction between word frequency and neighborhood size was marginally significant, $F_s(1, 39) = 3.90$, $p = .055$, $MSE = 366.80$; $F_i < 1$. For the low-frequency words, responses to words with large neighborhoods were an average of 10.5 ms faster than responses to words with small neighborhoods, $F_s(1, 39) = 10.18$, $p < .005$, $MSE = 467.27$; $F_i(1, 100) = 1.62$, $p = .206$, $MSE = 2,389.18$. For the high-frequency words there was no neighborhood size effect (both F's < 1).

The interaction between word frequency and neighborhood frequency was also significant, $F_s(1, 39) = 5.16$, $p < .05$, $MSE = 354.88$; $F_i < 1$. For the low-frequency words, responses to words with higher frequency neighbors were an average of 10.5 ms faster than responses to words with no higher frequency neighbors, $F_s(1, 39) = 8.11$,

Table 4

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 1B

Neighborhood frequency	Neighborhood size: Small	Large
	Low-frequency words	
No HFN	569 (5.5)	562 (4.5)
HFN	562 (5.1)	548 (3.9)
	High-frequency words	
No HFN	525 (1.8)	518 (1.5)
HFN	519 (1.1)	521 (1.2)

Note. HFN = higher frequency neighbors. Error rates in parenthesis ().

$p < .01$, MSE = 560.16; $F_i(1, 100) = 1.47$, $p = .229$, MSE = 2,389.18. For the high-frequency words there was no neighborhood frequency effect (both F's < 1). The neighborhood size by neighborhood frequency interaction was not significant (both F's < 1).

The three-way interaction between word frequency, neighborhood size, and neighborhood frequency was marginally significant, $F_s(1, 39) = 3.86$, $p = .057$, MSE = 358.59; $F_i < 1$. An inspection of Table 4 reveals that for the low-frequency words, words with large neighborhoods appeared to benefit more from the presence of higher frequency neighbors than did the words with small neighborhoods, whereas for the high-frequency words neither neighborhood size nor neighborhood frequency had any effect.

Error rates. The main effect of word frequency was significant, $F_s(1, 39) = 53.38$, $p < .001$, MSE = 16.96; $F_i(1, 200) = 28.98$, $p < .001$, MSE = 20.91. Participants made fewer errors to high-frequency words than to low-frequency words (1.4% vs. 4.7%). The main effect of neighborhood size and the main effect of neighborhood frequency were not significant ($F_s(1, 39) = 2.36$, $p = .132$, MSE = 13.22; $F_i(1, 200) = 1.29$, $p = .257$, MSE = 20.91; $F_s(1, 39) = 1.70$, $p = .2$, MSE = 10.90; $F_i < 1$, respectively), nor were any of the interactions (all F's < 1).

Discussion

The nonwords used in this experiment all had small neighborhoods. Because only the words with large neighborhoods would generate more global lexical activity than the nonwords, participants should have set their Σ criterion below the M criterion, and the multiple read-out model would predict that a facilitatory neighborhood size effect should be observed. Further, 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 some responses would be based upon the M criterion being exceeded, in which case an inhibitory neighborhood frequency effect should have been observed.

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). The neighborhood frequency effect, however, was facilitatory, not inhibitory, which is the opposite of what the model would predict.

Two additional results are of note. First, unlike the situation witnessed in Experiment 1A, 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 no effect on response latencies or errors. Second, the word frequency effect was larger in this experiment than in Experiment 1A (39.5 ms vs. 24.5 ms), as indicated by a significant interaction between experiment and word frequency, $F(1, 78) = 12.95$, $p < .005$, $MSE = 700.04$. This confirms that the word/nonword discriminations were more difficult when the nonwords had small neighborhoods than when they had no neighbors.

Experiment 1C: Lexical Decision with Large Neighborhood Nonwords

Method

Participants, stimuli, design, and procedure. The participants, word stimuli, design, and procedure are described in the General Method. The nonword stimuli consisted of 104 four-letter and 104 five-letter, orthographically-legal and pronounceable letter strings. The nonwords all had large neighborhoods (range of 6 to 21 neighbors, with a mean neighborhood size of 9.8).

Results

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. A total of 13 observations (0.16% of the data) were removed by this procedure. The mean response latencies of correct responses and the mean error rates are shown in Table 5. The mean response latency and percent error for the nonwords were 670 ms and 4.5%, respectively.

Response latencies. There was a significant main effect of word frequency, $F_s(1, 39) = 159.64$, $p < .001$, MSE = 983.0; $F_i(1, 200) = 60.75$, $p < .001$, MSE = 1,868.68. Responses to high-frequency words were an average of 44.5 ms faster than responses to low-frequency words. The main effect of neighborhood frequency was also significant, $F_s(1, 39) = 5.60$, $p < .05$, MSE = 504.88; $F_i(1, 200) = 1.38$, $p = .242$, MSE = 1,868.68. Responses to words with higher frequency neighbors were an average of 6 ms faster than responses to words without higher frequency neighbors. The main effect of neighborhood size was not significant, $F_s(1, 39) = 2.72$, $p = .107$, MSE = 571.09; $F_i < 1$, nor were any of the interactions (all p's > .10).

But an inspection of Table 5 reveals that, for the high-frequency words, there appeared to be an inhibitory neighborhood size effect. This observation was 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, $F_s(1,39) = 5.48$, $p < .05$, MSE = 643.54; $F_i(1, 100) =$

Table 5

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 1C

	Neighborhood size	
Neighborhood frequency	Small	Large
	Low-frequency words	
No HFN	609 (4.5)	607 (3.8)
HFN	599 (3.7)	600 (4.1)
	High-frequency words	
No HFN	557 (2.3)	565 (2.8)
HFN	552 (0.9)	563 (2.1)

Note. HFN = higher frequency neighbors. Error rates in parenthesis ().

2.59, $p = .111$, $MSE = 911.79$. For the low-frequency words there was no effect of neighborhood size ($F < 1$).

Error rates. The main effect of word frequency was significant, $F_s(1, 39) = 26.78$, $p < .001$, $MSE = 11.88$; $F_i(1, 200) = 10.07$, $p < .005$, $MSE = 20.07$. Participants made fewer errors to high-frequency words than to low-frequency words (2.0% vs. 4.0%). The main effect of neighborhood frequency was marginally significant, $F_s(1, 39) = 3.69$, $p = .062$, $MSE = 9.14$; $F_i(1, 200) = 1.17$, $p = .280$, $MSE = 20.07$. Participants generally 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 ($F_s < 1$;$F_i < 1$), nor were any of the interactions (all p'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. 1.6%, respectively), $F_s(1,39) = 3.44$, $p = .071$, $MSE = 8.71$; $F_i(1, 100) = 2.35$, $p = .128$, $MSE = 9.21$, 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 ($F < 1$).

Discussion

The nonwords employed in this experiment all 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 so that the Σ criterion would generally be set higher than the M criterion. Thus, the model predicts that most of the responses should be based on the M criterion, and an inhibitory neighborhood frequency effect should be observed. The model also predicts a null or small facilitatory neighborhood size effect due to the occasional use of the Σ criterion for responding.

The important findings are as follows. First, 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, just the opposite was true, as words with higher frequency neighbors were responded to more rapidly than words without higher frequency neighbors.

Second, there was some evidence of an inhibitory neighborhood size effect for the high-frequency words in this experiment. That is, responses 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 been facilitatory (Sears et al., 1995, Experiment 1). This issue will be addressed again in Experiment 3, which was an attempt to replicate the neighborhood size effects observed in this experiment (i.e., no effect of neighborhood size for low-frequency words and an inhibitory effect of neighborhood size for the high-frequency words), as these effects will figure prominently in the efforts to explain the neighborhood size and neighborhood frequency effects witnessed in Experiments 1A-1D.

Finally, note that the word frequency effect was larger in this experiment (44.5 ms) than in Experiment 1A (24.5 ms), $\underline{F}(1, 78) = 16.10$, $\underline{p} < .001$, $\underline{MSE} = 868.46$, and was approximately equivalent to that observed in Experiment 1B (39.5 ms), $\underline{F} < 1$. This would suggest that the word/nonword discriminations in this experiment were more difficult than they were in Experiment 1A, and were of similar difficulty to those in

Experiment 1B.

Experiment 1D: Lexical Decision with Small and Large Neighborhood Nonwords

Method

Participants, stimuli, design, and procedure. The participants, word stimuli, design, and procedure are described in the General Method. The nonword stimuli consisted of 104 four-letter and 104 five-letter, orthographically-legal and pronounceable letter strings. Half of the nonwords had small neighborhoods (range of 1 to 5 neighbors, with a mean neighborhood size of 3.3; this was a subset of the small neighborhood nonwords used in Experiment 1B), and half of the nonwords had large neighborhoods (range of 6 to 18 neighbors, with a mean neighborhood size of 9.8; this was a subset of the large neighborhood nonwords used in Experiment 1C). The overall mean neighborhood size of the nonwords was 6.5. As noted, the nonwords were matched closely to the words on neighborhood size. Specifically, the mean neighborhood size of the small neighborhood words and nonwords were 3.3 and 3.4, respectively, and the mean neighborhood size of the large neighborhood words and nonwords were both 9.8.

Results

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. A total of 28 observations (0.33% of the data) were removed by this procedure. The mean response latencies of correct responses and the mean error rates are shown in Table 6.

The mean response latency and percent error for the small neighborhood nonwords were 634 ms and 3.3%, respectively. The mean response latency and percent

Table 6

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 1D

Neighborhood frequency	Neighborhood size: Small	Large
	Low-frequency words	
No HFN	611 (5.4)	593 (4.0)
HFN	594 (3.3)	581 (2.8)
	High-frequency words	
No HFN	545 (2.0)	541 (2.0)
HFN	542 (0.9)	544 (1.0)

Note. HFN = higher frequency neighbors. Error rates in parenthesis ().

error for the large neighborhood nonwords were 664 ms and 5.4%, respectively. Responses to nonwords with large neighborhoods were an average of 30 ms slower than responses to nonwords with small neighborhoods, $F_s(1, 39) = 104.15$, $p < .001$, $MSE =$ 175.08. Participants also made more errors to nonwords with large neighborhoods than to nonwords with small neighborhoods (5.4% vs. 3.3%, respectively), $F_s(1, 39) = 30.44$, $p < .001$, $MSE = 3.03$.

Response latencies. In the analysis of the word data, there was a significant main effect of word frequency, $F_s(1, 39) = 436.66$, $p < .001$, $MSE = 490.20$; $F_i(1, 200) = 79.04$, $p < .001$, $MSE = 1{,}864.85$. Responses to high-frequency words were an average of 51.5 ms faster than responses to low-frequency words. The main effect of neighborhood size and the main effect of neighborhood frequency were significant, $F_s(1, 39) = 6.68$, $p < .05$, $MSE = 761.04$; $F_i(1, 200) = 1.59$, $p = .209$, $MSE = 1{,}864.85$, and $F_s(1, 39) = 7.71$, $p < .01$, $MSE = 585.87$, $F_i(1, 200) = 1.97$, $p = .162$, $MSE = 1{,}864.85$, respectively. 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.

There was a significant interaction between word frequency and neighborhood size, $F_s(1, 39) = 9.38$, $p < .005$, $MSE = 453.86$; $F_i(1, 200) = 1.54$, $p = .216$, $MSE =$ 1,864.85. For the low-frequency words, responses to words with large neighborhoods were an average of 15.5 ms faster than responses to words with small neighborhoods ($F_s(1, 39) = 13.31$, $p < .005$, $MSE = 700.48$; $F_i(1, 100) = 2.07$, $p = .153$, $MSE =$ 2,812.10). For the high-frequency words there was no neighborhood size effect (both

F's < 1). There was also a significant interaction between word frequency and neighborhood frequency, $F_s(1, 39) = 4.25$, $p < .05$, MSE = 940.64; $F_i(1, 200) = 1.78$, $p = .184$, MSE = 1,864.85. 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, $F_s(1, 39) = 7.78$, $p < .01$, MSE = 1,093.09; $F_i(1, 100) = 2.49$, $p = .118$, MSE = 2,812.10. For the high-frequency words there was no neighborhood frequency effect (both F's < 1). No other interactions were significant (all p's > .20).

Error rates. The main effect of word frequency was significant, $F_s(1, 39) = 44.02$, $p < .001$, MSE = 10.71; $F_i(1, 200) = 18.74$, $p < .001$, MSE = 17.01, as participants made fewer errors to high-frequency words than to low-frequency words (1.4% vs. 3.8%). The main effect of neighborhood frequency was also significant, $F_s(1, 39) = 12.01$, $p < .005$, MSE = 11.63; $F_i(1, 100) = 5.34$, $p < .05$, MSE = 17.01). Fewer errors were made to words with higher frequency neighbors than to words with no higher frequency neighbors (2.0% vs. 3.3%). The main effect of neighborhood size was not significant, $F_s(1, 39) = 2.10$, $p = .156$, MSE = 7.96; $F_i < 1$, nor were any of the interactions (all p's > .10).

Discussion

In this experiment 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 in these conditions. 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 neighborhood frequency were facilitatory for low-frequency words. (There was no effect of neighborhood size nor of neighborhood frequency for the high-frequency words).

Finally, note that the word frequency effect in this experiment (51.5 ms) was larger than that observed in Experiment 1A (24.5 ms), $\underline{F}(1, 78) = 53.06$, $\underline{p} < .001$, $\underline{MSE} = 556.28$, Experiment 1B (39.5 ms), $\underline{F}(1, 78) = 9.25$, $\underline{p} < .005$, $\underline{MSE} = 633.96$, and Experiment 1C (44.5 ms), $\underline{F}(1, 78) = 3.00$, $\underline{p} = .087$, $\underline{MSE} = 736.60$. These results suggest that the word/nonword discriminations were the most difficult in this experiment, as would be predicted by the multiple read-out model.

Summary of Experiments 1A-1D

Considered 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 only occur 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 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 1C the nonwords had large neighborhoods, and the model's prediction of a null neighborhood size effect was upheld.

The problem for the model is the facilitatory neighborhood size effect observed in Experiment 1D, where 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 no neighborhood size effect should occur, when in fact the largest facilitatory neighborhood size effect of all of the experiments was observed. Indeed, the model would predict 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 1D (9.5 ms vs. 15.5 ms, respectively). Overall then, the model does not appear to provide a viable 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 nonwords had large neighborhoods, and in Experiment 1D the words and the nonwords were matched on neighborhood size, and 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.[6] Although these results are quite consistent with those reported by Andrews (1989, 1992), Forster and Shen (1996), and Sears et al. (1995), they are directly the opposite of what the multiple read-out model would predict.

Experiment 2

As noted, 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. As noted, in those experiments, because the nonwords had large neighborhoods (Experiment 1C) or were matched to the words on neighborhood size (Experiment 1D), most responses should have been based on the M criterion, in which case an inhibitory neighborhood frequency effect should have been observed. Instead, the effect of higher frequency neighbors was facilitatory, not inhibitory, a result which is very difficult for the model to accommodate.

Because these results are so problematic for the model, 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 use the Σ criterion for responding. More specifically, in Experiment 2A the words and the nonwords all had small neighborhoods, and the words had either no higher frequency neighbors or exactly one higher frequency neighbor. Similarly, in Experiment 2B the words and the nonwords all 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 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 higher than the M criterion, and all responses should be made using the M criterion, in which case words with one higher frequency neighbor should be responded to more slowly than words without a higher frequency neighbor.

Method

Participants. Eighty undergraduate students from the University of Calgary participated in the experiment. Forty participated in Experiment 2A (words and nonwords with small neighborhoods), and forty 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 Appendix B, and the descriptive statistics for these stimuli are listed in Table 7).

In each experiment, half the stimuli were four-letter words, and the other half of the stimuli were five-letter words. In Experiment 2A all of the words were of low-

Table 7

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

	No HFN	One HFN
Experiment 2A		
Frequency	19.1	20.1
N	3.3	3.5
NBF	8.0	301.0
Experiment 2B		
Frequency	23.8	23.1
N	8.7	9.4
NBF	13.6	290.1

Note. HFN = higher frequency neighbor(s). NBF refers to the average frequency of the highest frequency neighbor.

frequency (mean Kucera and Francis normative frequency per million words of 19.6, range of 1 to 49), and had small neighborhoods (range 1 to 5 neighbors, with a mean neighborhood size of 3.4). In Experiment 2B all of the words were of low-frequency (mean Kucera and Francis normative frequency per million words of 22.9, range of 1 to 48), and had large neighborhoods (range of 6 to 17 neighbors, with a mean neighborhood size of 9.1). Approximately two-thirds of the words used in these Experiments were used in Experiments 1A-1D.

In each experiment the single factor manipulated was neighborhood frequency: The words had either no neighbors of higher frequency than themselves, or exactly one neighbor of higher frequency. For the words with a higher frequency neighbor, the mean Kucera and Francis 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.

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 of 1 to 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 of 6 to 18), with a mean neighborhood size of 9.1. These nonwords were a subset of those used in Experiment 1C.

<u>Design.</u> In each experiment the two neighborhood frequency conditions (no higher frequency neighbors, 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 carried out.

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 (with and without a higher frequency neighbor) and 15 orthographically-legal and pronounceable nonwords with small neighborhoods (i.e., at least one and no more than 5 neighbors). In Experiment 2B the practice trials consisted of 15 low-frequency words (with and without higher frequency neighbors) and 15 orthographically-legal and pronounceable nonwords with large neighborhoods (i.e., with at least 6 neighbors). (These practice stimuli were not used in the experiments proper, 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 (percent error) 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

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. In Experiment 2A a total of 4 observations (0.17% of the data) were removed by this procedure, and in Experiment 2B a total of 18 observations (0.50% of the data) were removed by this procedure. The mean response latencies of correct responses and the mean error rates for each experiment are shown in Table 8. The mean response latency and percent error for the small neighborhood nonwords used in Experiment 2A were 608 ms and 4.6%, respectively. The mean response latency and percent error for the large neighborhood nonwords used in Experiment 2B were 668 ms and 6.4%, respectively.[7]

Experiment 2A: Words and nonwords with small neighborhoods. In the analysis of the response latencies, the effect of neighborhood frequency was significant, $F_s(1, 39) = 12.48$, $p < .005$, MSE = 273.76; $F_i < 1$. Responses to words with one higher frequency neighbor were an average of 13 ms faster than responses to words without a higher frequency neighbor. 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, $F_s(1, 39) = 22.21$, $p < .001$, MSE = 305.49; $F_i < 1$. Responses to words with one higher frequency neighbor were an average of 19 ms faster than responses to words without a higher frequency neighbor. In the analysis of the error data, the effect of neighborhood frequency was not significant (both F's < 1).

Discussion

In these experiments the nonwords were matched to the words on neighborhood size. As was the case in Experiment 1D, 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

Table 8

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 2A (Words and Nonwords with Small Neighborhoods) and Experiment 2B (Words and Nonwords with Large Neighborhoods)

	No HFN	One HFN
Experiment 2A	556 (4.0)	543 (4.1)
Experiment 2B	605 (3.7)	586 (3.6)

Note. HFN = higher frequency neighbor(s).

(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 1D, the neighborhood frequency effect was facilitatory, not inhibitory.

Experiment 3

Thus far, the results of this investigation are fairly straightforward. For low-frequency words, the effect of higher frequency neighbors was consistently facilitatory in the previous experiments, even when the word/nonword context was manipulated to make the discrimination between the words and nonwords as difficult as possible (Experiment 1D and Experiment 2). With respect to the effect of neighborhood size, for low-frequency words, large neighborhoods would seem to facilitate responses in a variety of nonword contexts. Specifically, when the nonwords have no neighbors (Experiment 1A), when they have small neighborhoods (Experiment 1B), and when the words and the nonwords are matched on neighborhood size (Experiment 1D), responses to words with large neighborhoods are faster and generally less error prone than responses to words with small neighborhoods. The notable exception, of course, were the results from Experiment 1C, where the nonwords had large neighborhoods.

Recall that in that experiment there was no effect of neighborhood size for the low-frequency words, and some evidence of an inhibitory neighborhood size effect for the high-frequency words. The purpose of Experiment 3 was to attempt to replicate both of these findings prior to any attempts to explain the effects of nonword context on the orthographic neighborhood effects witnessed in the previous experiments.

To create an optimal situation to observe these effects, a more extreme neighborhood size manipulation was adopted relative to that used in Experiment 1C. Specifically, in Experiment 3 words had either a large neighborhood or no neighbors. As in Experiment 1C, the nonwords all had large neighborhoods. In Experiment 3A high-frequency words were used, and in Experiment 3B low-frequency words were used.

Method

Participants. Eighty undergraduate students from the University of Calgary participated in the experiment: forty participants in Experiment 3A and forty participants in Experiment 3B. 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. The stimuli for Experiment 3A consisted of five-letter high-frequency words, and the stimuli for Experiment 3B consisted of five-letter low-frequency words. The single factor manipulated in each experiment was neighborhood size. In Experiment 3A half of the words had large neighborhoods (range 6 to 15 neighbors), with a mean neighborhood size of 8.5, and the remainder of the words had no neighbors. For the words with large neighborhoods the mean number of higher frequency neighbors was 0.6 (range of 0 to 5), and the mean Kucera and Francis normative frequency of the highest frequency neighbor was 198.7. There were 20 words in each of the two conditions, for a total of 40 words.

In Experiment 3B half of the words had large neighborhoods (range of 6 to 15 neighbors), with a mean neighborhood size of 9.8, and the remainder of the words had no neighbors. For the words with large neighborhoods the mean number of higher frequency neighbors was 1.6 (range of 1 to 7), and the mean Kucera and Francis normative frequency of the highest frequency neighbor was 262.2. There were 30 words in each of the two conditions, for a total of 60 words. (The complete set of experimental words used in Experiments 3A and 3B is presented in Appendix C, and the descriptive statistics for these stimuli are listed in Table 9).

Two sets of nonwords with large neighborhoods were created. In Experiment 3A the mean neighborhood sizes of the large neighborhood words and the large neighborhood nonwords were identical (8.5; range of 6 to 14). In Experiment 3B the mean neighborhood sizes of the large neighborhood words and the large neighborhood nonwords were almost identical (9.8 and 9.7, respectively; range of 6 to 17). All of the nonwords were orthographically-legal and pronounceable.

Design. In each experiment the two neighborhood size conditions (words with no neighbors, words with large neighborhoods) 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 carried out.

Apparatus and procedure. The apparatus and procedure were identical to those used in Experiment 2.

In each experiment participants completed 20 practice trials prior to the collection of data. In Experiment 3A the practice trials consisted of 10 high-frequency words (five with large neighborhoods, five with no neighbors) and 10 nonwords with large neighborhoods. In Experiment 3B the practice trials consisted of 10 low-frequency words (five with large neighborhoods, five with no neighbors) and 10 nonwords with large neighborhoods. (These practice stimuli were not used in the experiment proper, and

Table 9

Mean Word Frequency and Neighborhood Size (N) for the Word Stimuli Used in Experiment 3A (High-Frequency Words) and Experiment 3B (Low-Frequency Words)

	Zero N	Large N
Experiment 3A		
Frequency	108.9	108.3
N	0.0	8.5
Experiment 3B		
Frequency	18.7	18.2
N	0.0	9.8

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 (percent error) of their responses, and during the experimental trials this information was presented every 20 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 the experiments was randomized separately for each participant.

Results

Response latencies less than 250 ms or more than 1,500 ms were considered outliers and were removed from the data set. Using this procedure, a total of 4 observations (0.25% of the data) were removed in Experiment 3A, and a total of 9 observations (0.37% of the data) were removed in Experiment 3B. The mean response latencies of correct responses and the mean error rates for Experiments 3A and 3B are shown in Table 10. The mean response latency and percent error for the large neighborhood nonwords in Experiment 3A were 642 ms and 3.7%, respectively. The mean response latency and percent error for the large neighborhood nonwords in Experiment 3B were 713 ms and 6.4%, respectively.[8]

Experiment 3A: High-frequency words. In the analysis of the response latencies, the main effect of neighborhood size was significant, $F_s(1, 39) = 5.16$, $p < .05$, $MSE = 492.85$; $F_i < 1$. Words with large neighborhoods were responded to an average of 12 ms more slowly than words with small neighborhoods. In the analysis of the error data, the main effect of neighborhood size was marginally significant, $F_s(1, 39) = 3.48$, $p = .070$, $MSE = 12.95$; $F_i < 1$. More errors were made to words with large neighborhoods than to words with small neighborhoods.

Table 10

Mean Response Latencies (in Milliseconds) and Error Rates (in %) in Experiment 3A (High-Frequency Words) and Experiment 3B (Low-Frequency Words)

	Zero N	Large N
Experiment 3A	549 (2.3)	561 (3.8)
Experiment 3B	606 (3.6)	609 (4.6)

Note. N = neighborhood size.

Experiment 3B: Low-frequency words. There was no effect of neighborhood size in the analysis of response latencies (both F's < 1), or in the analysis of error rates, $F_s(1, 39) = 2.18$, $p = .148$, MSE = 9.17; $F_i < 1$.

Discussion

Recall that in Experiment 1C (nonwords with large neighborhoods) there was no effect of neighborhood size for the low-frequency words, and some evidence of an inhibitory neighborhood size effect for the high-frequency words. The purpose of Experiment 3 was to attempt to replicate these findings. As noted, to create an optimal situation to observe these effects, a more extreme neighborhood size manipulation was employed relative to that used in Experiment 1C. Specifically, in Experiment 3 words had either a large neighborhood or no neighbors. As in Experiment 1C, the nonwords all had large neighborhoods.

In Experiment 3A, where only high-frequency words were used, an inhibitory neighborhood size effect was observed. That is, responses to words with large neighborhoods were slower and more error prone than responses to words with small neighborhoods. This result thus replicates that of Experiment 1C, where an inhibitory neighborhood size effect was also present in both the response latency and the error data for the high-frequency words. In Experiment 3B, where only low-frequency words were used, there was no effect of neighborhood size. This result also replicates that of Experiment 1C.

Having now established that the inhibitory neighborhood size effect for the high-frequency words in Experiment 1C was not a spurious result, the explanation for this effect (as well as the null effect of neighborhood size for the low-frequency words) will be addressed in the General Discussion after the multiple read-out model's ability to accommodate the results of this entire investigation is evaluated.

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' (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. That is, in Experiment 1A the nonwords had no neighbors, in Experiment 1B the nonwords had small neighborhoods, in Experiment 1C the nonwords had large neighborhoods, and in Experiment 1D the nonwords 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.

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 this 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 an inhibitory neighborhood frequency effect, and although the neighborhood size effect was facilitatory, the effect of neighborhood frequency was also facilitatory, not inhibitory (both effects interacted with word frequency, as they were observed only for the low-frequency words). In Experiment 1C the model predicted a facilitatory (or null) neighborhood size effect and an inhibitory neighborhood frequency effect, and yet neither of these effects were fully realized. That is, for the low-frequency words there was no effect of neighborhood size, but 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. In Experiment 1D the model predicted no effect of neighborhood size and an inhibitory neighborhood frequency effect, but both effects were facilitatory for the low-frequency words (there was no effect of neighborhood size nor neighborhood frequency for the high-frequency words).

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 would predict that an inhibitory neighborhood frequency effect should be observed. However, the neighborhood frequency effect was again facilitatory, both for words with small neighborhoods and 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 1D, where the words and the nonwords were matched on neighborhood size. According to the model, because 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) in this situation, no neighborhood size effect should be observed. That is, because the model assumes that a facilitatory neighborhood size effect will only occur when participants base their responses on global lexical activity (i.e., when they use the Σ criterion for responding), when participants cannot use the Σ criterion for responding facilitatory neighborhood size effects should not be observed. Thus, the facilitatory neighborhood size effect witnessed in this experiment casts serious doubt on the model's assumption that neighborhood size effects are due to a non-lexical mechanism (i.e., 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., intra-level 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 1C, the nonwords had large neighborhoods, and in Experiment 1D and in Experiment 2 the words and the nonwords were matched on neighborhood size, and yet in all of these experiments facilitatory neighborhood frequency effects were observed. These results cast serious doubt on the model's assumption that the inhibitory neighborhood frequency effect is a lexical selection effect intrinsic to word recognition.

<u>Accounting for Orthographic Neighborhood Effects in the Present Experiments</u>

Given that the multiple read-out model cannot accommodate the results of the present experiments, it is necessary to ask whether the model could do so if some of its assumptions were changed. The model could in fact accommodate most of the findings of the present investigation if one rather critical change was made; namely, that large neighborhoods and higher frequency neighbors facilitate word identification via both lexical and non-lexical selection mechanisms.

In the current model inhibition between lexical units is assumed to play a major role during the process of word recognition. That is, the lexical units of the higher frequency neighbors of the presented word are assumed to inhibit the word's lexical unit, delaying its accumulation of activation necessary to reach a critical activation threshold (the M criterion). This lexical inhibition mechanism embodied in the model is used to explain inhibitory neighborhood frequency effects, and because facilitatory neighborhood size effects cannot be explained via such a mechanism (as demonstrated in simulations conducted by Jacobs & Grainger, 1992), the model assumes that such effects are not genuine lexical selection effects (i.e., that they are due solely to the use of the Σ criterion).

But what if the role of lexical inhibition in the model was substantially reduced or even eliminated? Could the model then account for facilitatory neighborhood size and facilitatory neighborhood frequency effects even when responses are based on lexical selection?[9]

Andrews (1989) has suggested that facilitatory effects of neighborhood size can be explained by the reciprocal feedback loop between the lexical and sublexical units in the interactive-activation model (on which the multiple read-out model is based) in the following manner. When a low-frequency word with a large neighborhood is presented to the model, its lexical representation and the representations of its neighbors will be activated. The activated lexical units of the word and its neighbors send top-down activation to the sublexical units, which in turn send activation back up to the lexical units of the word and its neighbors. This reciprocal activation continues until the lexical unit corresponding to the target word has reached its activation threshold at which point lexical selection is achieved. A low-frequency word with a small neighborhood would not benefit as much from the reciprocal activation process because it has fewer neighbors contributing to the process. The end result would be that low-frequency words with large neighborhoods would reach a threshold activation level necessary for word identification before low-frequency words with small neighborhoods.

For high-frequency words the explanation is somewhat different. Because high-frequency words have higher resting activation levels than low-frequency words in the interactive-activation model, the assumption is that these words reach their activation threshold sufficiently rapidly through direct activation alone (Andrews, 1989). High-frequency words therefore do not benefit from the reciprocal activation as low-frequency words do, and thus there should be no effect of either neighborhood size nor neighborhood frequency for high-frequency words (as is most commonly observed in the literature; e.g., Andrews, 1989, 1992; Sears et al., 1995).

The interactive-activation model may also be capable of producing a facilitatory neighborhood frequency effect. As Sears et al. (1995) have suggested, if the amount of activation a neighbor sends back down to the sublexical units is related to its resting activation level, then higher frequency neighbors should send stronger top-down activation, due to their higher resting activation levels, than lower frequency neighbors. This would thus facilitate the speed with which the lexical unit corresponding to a low-frequency word reaches its activation threshold. Low-frequency words without higher frequency neighbors, of course, would not benefit from this increase in activation. (Again, because high-frequency words could reach an activation threshold through direct activation alone, they would not be affected by this top-down activation from higher frequency neighbors). Thus, the end result would be that low-frequency words with higher frequency neighbors would reach a threshold activation level necessary for word identification before low-frequency words without higher frequency neighbors.

As noted, the model currently assumes that when responses are based on global lexical activation, words with large neighborhoods and words with higher frequency neighbors will exceed the Σ criterion faster than words with small neighborhoods or words without higher frequency neighbors, and thus both the neighborhood size effect and the neighborhood frequency effect should be facilitatory. If the model also assumed that these effects would be facilitatory when responses are based on lexical selection (i.e., when the M criterion is exceeded), then a number of findings of the present investigation become explicable to the model. That is, the multiple read-out model, modified in this way, could account for the present results as follows.

In Experiment 1A the nonwords had no neighbors, and would therefore produce little lexical activity. In this situation, the Σ criterion can be set lower than the M criterion, because the lexical activity that the nonwords produce should overlap very little (if at all) with the lexical activity produced by the words. Thus, most responses could be made prior to lexical selection. A facilitatory neighborhood size effect and a facilitatory neighborhood frequency effect should then occur because words with large neighborhoods and words with higher frequency neighbors should, on average, produce more lexical activity than words with small neighborhoods and words with no higher frequency neighbors. In addition, because most responses would be based on lexical activation instead of lexical selection, the word frequency effect should be fairly small. This is what was observed in this experiment.

Recall that a facilitatory neighborhood size effect was also observed for the high-frequency words in this experiment. This result would also be expected, 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. The neighborhood frequency effect, on the other hand, was limited to the low-frequency words, perhaps because the number of neighbors is more strongly correlated with lexical activity than is the existence of higher frequency neighbors, particularly for high-frequency words (see Grainger & Jacobs, 1996, for a similar view).

In Experiment 1B the nonwords had small neighborhoods. First, consider the predictions for the low-frequency words. Under these circumstances only the words with large neighborhoods would have produced enough lexical activity to be distinguishable from the nonwords with small neighborhoods (the words with small neighborhoods and the nonwords would have produced comparable levels of lexical activity). Consequently, most of the words with large neighborhoods would be responded to prior to lexical selection (i.e., via the Σ criterion), whereas most of the words with small neighborhoods would be responded to following lexical selection. Responses to words with large neighborhoods would thus be faster than responses to words with small neighborhoods (a facilitatory neighborhood size effect), which was in fact what was observed.

The neighborhood frequency effect would also be expected to be facilitatory, but for slightly more complicated reasons. For the words with small neighborhoods, because they would generally be responded to following lexical selection, and because higher frequency neighbors would facilitate the lexical selection process (via the reciprocal activation mechanism embodied in the model), words with higher frequency neighbors should be responded to more rapidly than words without higher frequency neighbors, as was observed. On the other hand, the words with large neighborhoods would generally be responded to prior to lexical selection (via the Σ criterion), and because the large neighborhood words with higher frequency neighbors would generate more lexical activity than the large neighborhood words without higher frequency neighbors, a facilitatory neighborhood frequency effect should have occurred (as was observed).

For the high-frequency words, because many of the responses would be based on lexical selection, and because neither neighborhood size nor neighborhood frequency would affect the lexical selection process for high-frequency words, both of these effects should be greatly attenuated or even eliminated. Thus, the lack of a neighborhood size effect or a neighborhood frequency effect for the high-frequency words in this experiment is consistent with this explanation. Also note that, relative to Experiment 1A, because more responses would be based on lexical selection, the word frequency effect should have been larger in this experiment, which was in fact the case (24.5 ms vs. 39.5 ms).

In Experiment 1C the nonwords had large neighborhoods, and to explain the results of this experiment within this framework one has to assume that lexical activity <u>can</u> be a reliable cue as to the lexicality of a letter string under these circumstances (contrary to the multiple read-out model's assumption). More specifically, in these situations, words with small neighborhoods and words without higher frequency neighbors should, on average, produce less lexical activity than the nonwords, and thus some of these stimuli could still be responded to using the $\sum$ criterion (i.e., prior to lexical selection). That is, if a stimulus produced very little lexical activity, then participants would be biased to respond "word". Conversely, words with large neighborhoods and words with higher frequency neighbors should, on average, produce comparable levels of lexical activity as the nonwords, and thus should be responded to following lexical selection (via the M criterion).

For the low-frequency words, a bias to respond "word" when a stimulus produces little lexical activation would mean that, overall, responses to words with small neighborhoods should be faster than responses to words with large neighborhoods (an inhibitory neighborhood size effect). Because the words with large neighborhoods are responded to following lexical selection, however, and because the lexical selection process is facilitated by large neighborhoods, this tendency toward an inhibitory neighborhood size effect would be attenuated or even eliminated, as was observed in this experiment.

Similarly, with regard to the effect of neighborhood frequency, for the words with large neighborhoods the neighborhood frequency effect would be facilitatory, because words with higher frequency neighbors would be processed more rapidly than words without higher frequency neighbors via lexical selection. For the words with small neighborhoods the prediction is not as clear. On the one hand, the words without higher frequency neighbors should have produced less lexical activity than the words with higher frequency neighbors, and if participants were biased to respond "word" when a stimulus produced very little lexical activity, then, overall, the neighborhood frequency effect should be inhibitory (i.e., the words without higher frequency neighbors would produce less lexical activity than the words with higher frequency neighbors, and therefore should be responded to more rapidly). On the other hand, if many of the words with higher frequency neighbors tended to be processed via the lexical selection mechanism (because they produced too much lexical activity to be responded to via the Σ criterion), then the processing of these words would be facilitated. The net effect would be an attenuated, null, or even a slightly facilitatory neighborhood frequency effect (as was observed), depending upon the particular stimulus characteristics of the words in question.

For the high-frequency words, a bias to respond "word" when a stimulus produces little lexical activity would again mean that, overall, responses to words with small neighborhoods should be faster than responses to words with large neighborhoods (an inhibitory neighborhood frequency effect). That is, high-frequency words with small neighborhoods would produce less lexical activity than high-frequency words with large neighborhoods, and so many of the small neighborhood words would be responded to prior to lexical selection (via the Σ criterion), whereas many of the large neighborhood words would be responded to following lexical selection. But unlike the situation for the low-frequency words, this tendency toward an inhibitory neighborhood size effect would not be counteracted by the facilitated lexical processing of large neighborhood words. That is, because the lexical processing of high-frequency words is assumed to be insensitive to neighborhood size, the neighborhood size effect would be expected to be inhibitory, and there was indeed a trend towards such an effect in this experiment. (There would be no neighborhood frequency effect for the large neighborhood words for the same reason, whereas for the small neighborhood words the effect would be expected to be null or perhaps slightly inhibitory).

This account is fairly speculative, and hinges on the notion that participants can use low levels of lexical activity to make fast "word" responses (via the Σ criterion). But the results of Experiment 3 do lend support to this idea. Recall that the purpose of this experiment was to replicate the main findings of Experiment 1C; namely no effect of neighborhood size for the low-frequency words, and an inhibitory neighborhood size effect for the high-frequency words. For this purpose the neighborhood size manipulation was more extreme in Experiment 3 (the words had no neighbors or large neighborhoods), and the results did in fact replicate those of Experiment 1C. Thus, Experiment 3 provided independent evidence that participants can use low levels of lexical activity to make fast "word" responses, with the end result being an inhibitory neighborhood size effect for high-frequency words and no neighborhood size effect for low-frequency words.

In Experiment 1D the words and the nonwords were matched on neighborhood size. In these situations lexical activity is not a reliable cue as to the lexicality of a letter string, and the Σ criterion should be set higher than the M criterion, so that most (if not all) of the responses should be made following lexical selection. In this situation, assuming that both large neighborhoods and higher frequency neighbors facilitate lexical selection, the effect of neighborhood size and the effect of neighborhood frequency should be facilitatory, as was observed. Also note that the word frequency effect was significantly larger in this experiment than in any of the others, which would be expected if more responses were being made following lexical selection in this experiment than in Experiments 1A-1C (Becker, 1976; Forster, 1976; McClelland & Rumelhart, 1981; Paap et al., 1982; Stone & Van Orden, 1993; Whaley, 1978).

Experiment 2 was a partial replication of Experiment 1D, in that neighborhood frequency was manipulated and the words and the nonwords completely overlapped in neighborhood size. In this experiment, as was the case in Experiment 1D, lexical activity was not a reliable cue as to the lexicality of a letter string, and most words should have been responded to following lexical selection. Thus, a facilitatory neighborhood frequency effect would be expected, as was observed. This finding further supports the assumption that higher frequency neighbors help, not hinder, the identification of low-frequency words.

The results of this analysis suggest that it is possible to explain both the neighborhood size effects and the neighborhood frequency effects observed in these experiments within the framework of the multiple read-out model, and, in particular, how these effects were modulated by the nonword context. The extent to which the required modifications to the model's assumptions would harm its ability to account for other word recognition phenomenon would of course need to be evaluated.

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) 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 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.

Recently, Sears, Hino, and Lupker (in press) 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 which 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 model's 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. Further, the effect of neighborhood size and the effect of neighborhood frequency were negatively and independently related to the error scores.
Because lower error scores correspond to faster lexical decision and pronunciation latencies in these models, these models thus predict that large neighborhoods and higher frequency neighbors should facilitate responses to low-frequency words.

Conclusions

The purpose of this investigation 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 results of this study show that large neighborhoods and higher frequency neighbors facilitate responses to low-frequency words under 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. However, by reducing or eliminating the role of lexical inhibition in the model, and leaving the reciprocal feedback loop as the critical lexical selection process, the multiple read-out model would be able to account for the facilitatory orthographic neighborhood effects observed in this study.

Endnotes

[1]Orthographic neighborhood effects have also been investigated in tasks other than lexical decision (for a review, see Andrews, 1997). In the pronunciation task participants are required to pronounce letter strings as rapidly and accurately as possible, and the unanimous finding is that large neighborhoods and higher frequency neighbors facilitate responses, particularly to low-frequency words (Andrews, 1989, 1992; Carreiras et al., 1997; Grainger, 1990; Sears et al., 1995; Peereman & Content, 1995).

In the semantic categorization task participants have typically been asked to make binary animal/nonanimal classifications (e.g., respond yes to "Horse" and respond no to "Dream"), and the "no" responses are the responses of interest. Using this task Carreiras et al. (1997) and Forster and Shen (1996) have reported no effect of neighborhood size and either no effect or an inhibitory effect of neighborhood frequency. Sears, Lupker, and Hino (in press) have recently reported a facilitatory neighborhood size effect (but no effect of neighborhood frequency).

The effects of orthographic neighbors have also been studied in perceptual identification tasks. Grainger and colleagues (Carreiras, et al., 1997; Grainger and Jacobs, 1996; Grainger & Segui, 1990) have used the progressive demasking task, in which each trial is a succession of target word and mask ("#####") presentations. Over the course of the trial the duration of the target word presentation is increased and the duration of the mask presentation is decreased, which serves to gradually increase the visibility of the target word. Participants are required to stop the trial when they think they have identified the target word, and then provide their response. These investigators have reported that responses to words with higher frequency neighbors (and to a lesser extent words with large neighborhoods) are slower than responses to words without higher frequency neighbors. Snodgrass and Minzer (1993) have used a task in which words were presented in a series of increasing fragments, and participants either made successive identification responses or a single identification response. They reported that when multiple responses were allowed, large neighborhoods either had no effect or a slight facilitatory effect, but when only one response was allowed words with large neighborhoods were identified less frequently than words with small neighborhoods (they did not examine the effects of higher frequency neighbors). Finally, when words were presented very briefly with both a forward and backward mask, Grainger and Jacobs reported an inhibitory effect of higher frequency neighbors, whereas Sears, Lupker, and Hino (in press) reported the opposite – words with higher frequency neighbors were identified more frequently than words without higher frequency neighbors (the neighborhood size effect was also facilitatory).

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

[3]Several investigators have proposed that lexical decisions can be made prior to lexical selection (Balota & Chumbley, 1984; Johnson & Pugh, 1994; Seidenberg & McClelland, 1989). For example, Balota and Chumbley have proposed that lexical decisions can be made based upon a familiarity/meaningful dimension. Similarly,

Johnson and Pugh have proposed that when a letter string is presented, global lexical activity ("cohort information") is available immediately, and that lexical decisions can be made based upon this information.

[4]Many of the item analyses reported in the present paper were not statistically significant, but this should not be taken as a particularly important issue. Although Clark (1973) has argued that items, as well as participants, should be considered as a random factor in these types of analyses, it is seldom the case that 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 2). Consequently, as Wike and Church (1976) and others (Cohen, 1976; Keppel, 1976) 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 due to a greatly deflated alpha value). Any concerns readers might have about the generalizability of the results across items are addressed by the fact that several important findings of Experiments 1C and 1D were replicated in Experiments 2 and 3, using a largely different set of items. 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 were significant by item analyses.

[5]This finding is inconsistent with that of Sears et al.'s (1995) Experiment 5, who reported a facilitatory neighborhood size effect under identical nonword conditions. The most likely reason for the difference between the results of Sears et al.'s Experiment 5

and those of Experiment 1C is that error rates were larger in the Sears et al. 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 1C). This would suggest that the Sears et al. participants were making more responses based on global lexical activity than the participants in Experiment 1C (see Grainger & Jacobs, 1996, Experiment 1D). To address this possibility, Experiment 1C was repeated in which instructions emphasizing speed over accuracy were provided to a new group of participants. If the Sears et al. results were due to their participants making many of their responses based on global lexical activity rather than lexical selection, then by inducing the same type of responding a facilitatory neighborhood size effect should be observed. This was in fact the case. There was also a significant word frequency by neighborhood size interaction, reflecting the fact that large neighborhoods facilitated responses only to 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% vs. 9.0% in the Sears et al. experiment).

[6]Carreiras et al. (1997) have suggested that any 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, a post hoc analysis was conducted 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 1D were still statistically significant. Second, for the high-frequency words in Experiment 1C, 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-1C were attenuated in this reanalysis (presumably due to loss of statistical power), and in no case were they inhibitory. The facilitatory neighborhood frequency effect observed in Experiment 1D, 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 where the presence of phonologically inconsistent orthographic neighbors has been controlled.

[7]The Author Recognition Test (ART; Stanovich & West, 1989) was administered to each participant after they finished Experiment 2 (it was not administered to the participants in Experiments 1A-1D, because these experiments were either completed or near completion when it was suggested that the ART be used). The ART is a short questionnaire that is comprised of 86 names, 45 of which are names of popular writers of books, magazine articles, and/or newspaper columns. Participants were asked to read

each name on the list and to put a check beside each and every name they knew to be a popular writer. (They were instructed not to guess). The ART is a measure of exposure to print, and it was administered to determine whether or not orthographic neighborhood effects vary across varying levels of exposure to print. This was done in the following manner.

The participants' ART questionnaires were first scored. This was done by summing the total number of checks and then subtracting from this the number of incorrect checks. The scores were then partitioned into quartiles, and the scores from the lowest quartile (the lowest scores) and the scores from the highest quartile (the highest scores) were then submitted to a 2 (Group: low exposure, high exposure) x 2 (Neighborhood Frequency: no higher frequency neighbors, one higher frequency neighbor) mixed-model analysis of variance (ANOVA), with group as a between-subjects factor and neighborhood frequency as a within-subjects factor. In Experiment 2A the mean score for the low exposure group was 9.9 (range of 6-12) and the mean score for the high exposure group was 27.4 (range of 21-37). In Experiment 2B there were 11 participants in each group (due to a tie for the highest score in the low exposure group and for the lowest score in the high exposure group). The mean score for the low exposure group was 9.4 (range 4-12) and the mean score for the high exposure group was 25.4 (range 20-35). (Only subject analyses were performed).

In the analysis of the response latencies from Experiment 2A, the main effect of group was marginally significant, $F(1, 18) = 3.97$, $p = .062$, $MSE = 3{,}681.56$. The high exposure group responded an average of 43 ms faster than the low exposure group. The main effect of neighborhood frequency was also significant, $F(1, 18) = 12.44$, $p < .005$, .

MSE = 241.99. Responses to words with one higher frequency neighbor were an average of 18 ms faster than responses to words with no higher frequency neighbors. In addition, the group by neighborhood frequency interaction was marginally significant, $F(1, 18) = 3.46$, $p = .079$, MSE = 241.99. Responses to words with one higher frequency neighbor were 27 ms faster than responses to words with no higher frequency neighbors for the high exposure group and only 9 ms faster for the low exposure group. In the analysis of the error rates, the main effects of group, neighborhood size, and the interaction between these two factors were all not significant (all F's < 1). The error rates for the high exposure group and the low exposure group were virtually identical (4.8% and 4.6%, respectively).

The nonword data were submitted to a one-factor (Group: low exposure, high exposure) between-subjects analysis of variance (ANOVA). In the analysis of response latencies, the high exposure group responded 43 ms faster than the low exposure group, but this difference was not statistically significant, $F(1, 18) = 2.25$, $p = .151$, MSE = 4,015.33. In the analysis of the error rates, the high exposure group responded incorrectly 4.5% of the time and the low exposure group 3.6% of the time, but this difference was not statistically significant, $F < 1$.

In the analysis of the response latencies of Experiment 2B, the main effect of group was not significant, $F < 1$, but the main effect of neighborhood frequency was significant, $F(1, 20) = 12.08$, $p < .005$, MSE = 362.83. Responses to words with one higher frequency neighbor were an average of 20 ms faster than responses to words with no higher frequency neighbors. The group by neighborhood frequency interaction was not significant, $F < 1$. In the analysis of the error rates, the main effect of group was marginally significant, $F(1, 20) = 4.04$, $p = .058$, $MSE = 16.01$, and the main effect of neighborhood frequency and the group by neighborhood frequency interaction were not significant (all F's < 1). The high exposure group made fewer errors than the low exposure group (2.4% vs. 4.8%).

In the analysis of response latencies to the nonwords, the high exposure group responded 81 ms faster than the low exposure group, and this difference was statistically significant, $F(1, 20) = 7.44$, $p < .05$, $MSE = 4{,}905.04$. In the analysis of the error rates, the high exposure group responded incorrectly 4.4% of the time and the low exposure group 5.8% of the time, but this difference was not statistically significant, $F(1, 20) = 1.01$, $p = .327$, $MSE = 10.87$.

To summarize, the important findings are as follows. First, this analysis replicated the analyses reported in Experiment 2 in that the effect of neighborhood frequency was facilitatory in both Experiments 2A and 2B. Second, in Experiment 2A the marginal group by neighborhood frequency interaction suggests that the facilitatory neighborhood frequency effect was larger for those with more exposure to print. Although not statistically significant, in Experiment 2B this same pattern of results was observed, as for the high exposure group the facilitatory neighborhood frequency effect was 25 ms, whereas it was 14 ms for the low group. Third, in general the high exposure groups responded faster to the nonwords than did the low exposure groups (this effect was significant in Experiment 2B). Finally, it should be noted that these analyses were hampered by a reduced statistical power because there were few participants in each group (due to the quartile splits). The purpose of explicitly testing how differences in exposure to print affect orthographic neighborhood effects in lexical decision was, however, beyond the scope of the present investigation.

[8]As was the case in Experiment 2, the ART was administered to each participant after they completed the experiment. The scores were again partitioned into quartiles, and the scores from the lowest quartile (the lowest scores) and the scores from the highest quartile (the highest scores) were submitted to a 2 (Group: low exposure, high exposure) x 2 (Neighborhood Size: no neighbors, large neighborhoods) mixed-model analysis of variance (ANOVA), with group as a between-subjects factor and neighborhood size as a within-subjects factor. In Experiment 3A the mean score for the low exposure group was 7.6 (range of 6-10) and the mean score for the high exposure group was 26.0 (range of 21-36). In Experiment 3B there were 11 participants in each group (due to a tie for the highest score in the low exposure group and for the lowest score in the high exposure group). The mean score for the low exposure group was 7.5 (range 6-9) and the mean score for the high exposure group was 22.4 (range 17-36). (Again, only subject analyses were performed).

In the analysis of the response latencies from Experiment 3A, the main effect of group was not significant, $F < 1$, and the main effect of neighborhood size was marginally significant, $F(1, 18) = 3.48$, $p = .078$, $MSE = 677.33$. Responses to words with large neighborhoods were an average of 15 ms slower than responses to words with no neighbors. The group by neighborhood size interaction was not significant, $F < 1$. In the analysis of the error rates, the main effect of group was not significant, $F < 1$, as the error rates for the high group and the low group were virtually identical (3.5% and 3.0%, respectively). The main effect of neighborhood size was marginally significant, $F(1, 18) = 4.33$, $p = .052$, $MSE = 14.44$, as, on average, more errors were made to words with large neighborhoods (4.5%) than to words with no neighbors (2.0%). The group by neighborhood size interaction was not significant, $F < 1$.

The nonword data were submitted to a one-factor (Group: low exposure, high exposure) between-subjects analysis of variance (ANOVA). In the analysis of response latencies, the high exposure group responded 18 ms faster than the low exposure group, but this difference was not statistically significant, $F < 1$. In the analysis of the error rates, the high exposure group responded incorrectly 3.7% of the time and the low exposure group 3.2% of the time, but this difference was not statistically significant, $F < 1$.

In the analysis of the response latencies from Experiment 3B, the main effects of group and neighborhood size were not significant (all p's > .20). The group by neighborhood frequency interaction was marginally significant, $F(1, 20) = 3.65$, $p = .071$, $MSE = 441.81$. For the high exposure group, responses to words with large neighborhoods were an average of 20 ms slower than responses to words with no neighbors. For the low exposure group, responses to words with large neighborhoods were an average of 5 ms faster than responses to words with no neighbors. In the analysis of the error rates, the main effect of group was marginally significant, $F(1, 20) = 3.74$, $p = .067$, $MSE = 15.20$, as the high exposure group made fewer errors than the low exposure group (3.7% vs. 6.0%, respectively). The main effect of neighborhood frequency and the group by neighborhood frequency interaction were not significant (all F's < 1).

In the analysis of response latencies to the nonwords, the high exposure group responded 44 ms faster than the low exposure group, but this difference was not statistically significant, $F < 1$. In the analysis of the error rates, the high exposure group responded incorrectly 7.4% of the time and the low exposure group 7.5% of the time, $F < 1$.

To summarize, the important findings are as follows. First, in Experiment 3A high-frequency words with large neighborhoods were generally responded to more slowly than high-frequency words without neighbors, a finding that replicates that of Experiment 1C and Experiment 3. Second, in Experiment 3B the marginal interaction between group and neighborhood size suggests that the readers with high exposure to print may have been able to reliably use lexical activity in making their responses. That is, they may have responded to the words with large neighborhoods using the Σ criterion, as these words were responded to 20 ms slower than the words with no neighbors, whereas the readers with less exposure to print showed no evidence of this type of response strategy. The note of caution about a lack of statistical power in these analyses (endnote 7) applies here as well. However, considered together the results of these analyses suggest that more research should be conducted to delineate how exposure to print influences orthographic neighborhood effects.

[9]Simulations regarding the importance of the excitatory connections between the lexical units and sublexical units of the interactive-activation model have thus far produced mixed results. For example, Jacobs and Grainger's (1992) Simulation 4 failed to replicate Andrews' (1992) finding of a facilitatory neighborhood for low-frequency words. They did state, however, "if this result of Andrews can be successfully replicated in further experimentation, then it clearly presents a challenge to the IA model" (Jacobs & Grainger, 1992, p. 1181). The findings from the present experiments, along with those

of Forster and Shen (1996) and Sears et al. (1995), thus represents just such a challenge to Grainger and Jacobs' implementation of the interactive-activation model. Further simulation work is required to determine if parameter values can be found that successfully simulate facilitatory orthographic neighborhood effects, while allowing the model to maintain its ability to account for other word identification phenomenon. Coltheart and Rastle's (1994) simulation work is a step in this direction. They implemented the architecture of the interactive-activation model as the lexical decision component of their dual-route cascade model, and have found parameter settings that allow their model to simulate facilitatory neighborhood size effects. They reported that when they set the parameter value responsible for the excitatory connections between the lexical and sublexical units to zero, their model's ability to simulate the facilitatory effect of neighborhood size disappeared.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? Journal of Experimental Psychology: Learning, Memory, & 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, & Cognition, 18, 234-254.

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

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

Becker, C.A. (1976). Allocation of attention during visual word recognition. Journal of Experimental Psychology: Human Perception & Performance, 2, 556-566.

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

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). Hilldale, NJ: Erlbaum.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. Journal of Experimental Psychology: Human Perception & Performance, 20, 1197-1211.

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

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. Journal of Memory & 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.

Jacobs, A.M., & Grainger, J. (1992). Testing a semistochastic variant of the interactive activation model in different word recognition experiments. Journal of Experimental Psychology: Human Perception & Performance, 18, 1174-1188.

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.

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, & 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 & 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.

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

Sears, C.R., Hino, Y., & Lupker, S.J. (In press). Orthographic neighborhood effects in parallel distributed processing models. Canadian Journal of Experimental Psychology.

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

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.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. Reading Research Quarterly, 24, 402-433.

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

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.

Whaley, C.P. (1978). Word-nonword classification time. Journal of Verbal Learning and Verbal Behavior, 17, 143-154.

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 A

Items Used in Experiments 1A-1D

(Note that an * indicates that these items have at least one inconsistent orthographic neighbor and were excluded from the reanalysis of Experiments 1A-1D).

High-Frequency/Small Neighborhood/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, STYLE

High-Frequency/Small Neighborhood/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, YOUTH*

High-Frequency/Large Neighborhood/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, WATCH*

High-Frequency/Large Neighborhood/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, STORE

Low-Frequency/Small Neighborhood/No Higher Frequency Neighbors: CRIB, DEBT, DIRT, DRUG, FUSE*, GASP*, GLAD, HURT, INCH, LION*, MONK*, STUD, TUBE*, BLAME, BLAST, BOOST, BRICK, CRAWL, GLAZE, GLOOM, HARSH, PLEAD, SAUCE, SLAVE, SPRAY, STEEL

Low-Frequency/Small Neighborhood/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, YIELD

Low-Frequency/Large Neighborhood/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, WIPED

Low-Frequency/Large Neighborhood/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, TRACE

Appendix B

Items Used in Experiments 2A and 2B

Small Neighborhood/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, STEEL

Small Neighborhood/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, YIELD

Large Neighborhood/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, WIPED

Large Neighborhood/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, TRACE

Appendix C

Items Used in Experiments 3A and 3B

High-Frequency/No Neighbors: AGREE, BELOW, CLAIM, DOUBT, EARTH, EMPTY, EQUAL, EXIST, FAITH, FINAL, IMAGE, ISSUE, KNIFE, MUSIC, OFFER, RAISE, VISIT, VOICE, WAGON, WHEEL

High-Frequency/Large Neighborhood: BEACH, CARRY, CLASS, CROSS, GRASS, LOWER, MODEL, PARTY, PLANE, REACH, RIVER, SHARE, SIGHT, SOUND, SPITE, STAGE, STONE, STORE, SWEET, WATCH

Low-Frequency/No Neighbors: ALARM, ARROW, BROIL, BRUTE, CLERK, CLIFF, CRUDE, CRUMB, DEVIL, DIGIT, EXTRA, FALSE, FANCY, FLUID, FROZE, GUARD, GHOST, LAUGH, MAPLE, PROOF, PROUD, RANCH, RINSE, SALAD, SUGAR, THEFT, THIGH, THORN, TWIST, WIDTH

Low-Frequency/Large Neighborhood: BAKED, CROWN, FIRED, GRADE, GREED, GROWN, HONEY, LIVER, LOVER, MOUSE, PANTS, PEACH, POKER, PRIME, SCARE, SEEDS, SHADE, SHAKE, SHAVE, SMELL, SPARE, SPICE, STACK, STALE, STARE, STARS, STOOP, TIGHT, TOWER, WAGER