Perception & Psychophysics
1999, 61 (8), 1537-1554

Orthographic neighborhood effects in perceptual identification and semantic categorization tasks: A test of the multiple read-out model

CHRISTOPHER R. SEARS
University of Calgary, Calgary, Alberta, Canada

STEPHEN J. LUPKER
University of Western Ontario, London, Ontario, Canada

and

YASUSHI HINO
Chukyo University, Nagoya, Japan

How should a word's orthographic neighborhood affect perceptual identification and semantic categorization, both of which require a word to be uniquely identified? According to the multiple read-out model (Grainger & Jacobs, 1996), inhibitory neighborhood frequency effects should be observed in these types of tasks, and facilitatory neighborhood size effects should not be. In Experiments 1 and 2 (perceptual identification), these effects were examined as a function of stimulus visibility (i.e., high vs. low visibility) to provide as full a test as possible of the model's predictions. In the high-visibility conditions, words with large neighborhoods were reported less accurately than words with small neighborhoods, but there was no effect of neighborhood frequency (i.e., whether the word had a higher frequency neighbor). In the low-visibility conditions, low-frequency words with large neighborhoods and low-frequency words with higher frequency neighbors showed superior identification performance. In the semantic categorization task (Experiment 3), words with large neighborhoods were responded to more rapidly than words with small neighborhoods, but there was no effect of neighborhood frequency. These results are inconsistent with two of the basic premises of the multiple read-out model—namely, that facilitatory neighborhood size effects are due to a variable response criterion (the Σ criterion), rather than to lexical selection processes, and that the lexical selection processes themselves produce an inhibitory neighborhood frequency effect (via the M criterion). Instead, the present results, in conjunction with previous findings, suggest that large neighborhoods (and perhaps higher frequency neighbors) do aid lexical selection.

A considerable amount of recent research in visual word recognition has been directed at determining whether the speed with which a word is identified is affected by the extent to which that word is orthographically similar to other words (see Andrews, 1997, for a review). For example, many of these studies have centered on the question of whether and how identification time might be related to Coltheart's N (Coltheart, Davelaar, Jonasson, & Besner, 1977). This metric indexes the size of a word's *orthographic neighborhood*, which is defined as the set of words that can be created by changing one letter of the word while preserving letter positions. For example, the words *pike*, *pine*, *pole*, and *tile* are all orthographic neighbors of the word *pile*. By this definition, some words are orthographically unique and possess no neighbors (e.g., *idol*), whereas other words possess large orthographic neighborhoods (e.g., *rate*, with 21 neighbors). For most models of word recognition, the size of a word's orthographic neighborhood does have processing implications.

The reason that the size of a word's neighborhood has processing implications for most models of word recognition is that, according to these models, the lexical representations of a word's orthographic neighbors become activated while the word is being recognized. Once activated, these units then play a role in the lexical selection process. Consider, for example, models that incorporate a serial-search mechanism, such as the activation–verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982). In this model, an initial spread of activation through a network of sublexical and lexical units serves to isolate a set of lexical candidates that are consistent with the gross sensory characteristics of the input stimulus. A more de-

---

Portions of these data were presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles, November 1995. We thank Lester Krueger, Philip Allen, Leann Stadtlander, and Susan Graham for their very valuable comments. Correspondence concerning this article should be addressed to C. R. Sears, Department of Psychology, University of Calgary, 2500 University Drive, Calgary, AB, T2N 1N4 Canada (e-mail: sears@ucalgary.ca).

—*Accepted by previous editor, Myron L. Braunstein*

tailed serial verification process then checks or verifies each candidate item to determine whether it matches the sensory representation. The verification process is frequency ordered, in such a way that high-frequency words in the candidate set are checked before low-frequency words. Lexical selection is accomplished when a correct match is found. Because a word's orthographic neighbors are highly similar to the word itself, they would tend to be members of the candidate set. Consequently, increases in the size of a word's neighborhood should produce increases in the size of the candidate set which should in turn produce increases in the time required for lexical selection. Thus, according to the activation–verification model, there should be an inhibitory neighborhood size effect—longer response latencies for words with large neighborhoods than for words with small neighborhoods.

Coltheart et al. (1977) were the first to specifically examine the effects of neighborhood size. In their study, in which word frequency was controlled but not manipulated, manipulations of neighborhood size had no effect on the lexical decision latencies to words. More recently, however, Andrews (1989) reported data suggesting that not only did neighborhood size affect response latency, but that it interacted with word frequency. In particular, in Andrews's experiments, lexical decision and naming latencies to low-frequency words with large neighborhoods were shorter than the latencies to low-frequency words with small neighborhoods. For high-frequency words, neighborhood size had little or no effect on response latencies.

Subsequent research (Andrews, 1992; Forster & Shen, 1996; Sears, Hino, & Lupker, 1995) has generally confirmed this facilitatory neighborhood size effect for low-frequency words (a notable exception being Johnson & Pugh, 1994, who reported that neighborhood size effects are inhibitory when pronounceable nonwords are used in a lexical decision task, a finding that was not replicated by Sears et al.). What should be noted is that these reports of facilitatory neighborhood size effects are clearly at odds with serial-search models (e.g., Forster, 1976; Paap et al., 1982), because, as we have noted above, larger neighborhoods should produce larger candidate sets, which should in turn increase the amount of time required for a verification or comparison process.

Although some investigators have focused on the neighborhood size effect and its implications for models of lexical selection, studies by Grainger and colleagues (Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989; Grainger & Segui, 1990) suggest that the frequency of a word's orthographic neighbors relative to its own frequency (referred to as *neighborhood frequency*) is also an important factor in word recognition. 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.

Because there was no difference in the lexical decision latencies between the first two conditions, Grainger et al. (1989) concluded that neighborhood size per se had little or no effect. However, responses to words with one higher frequency neighbor were slower than responses to words with no higher frequency neighbors. Grainger et al. (1989) argued that, consistent with the account provided by serial-search models, this result indicates that the existence of higher frequency neighbors delays lexical selection. As Grainger et al. (1989) noted, serial-search models would predict that higher frequency neighbors of a target word would delay lexical selection, because a target word's higher frequency neighbors would have to be evaluated first during the frequency-ordered search for the target's representation. As Grainger et al. (1989) also noted, however, these models further predict a cumulative inhibitory neighborhood frequency effect, because as the number of higher frequency neighbors increases, so too would the delay in finding the target's lexical representation. This prediction was not upheld. That is, responses to words with many higher frequency neighbors were not significantly slower than responses to words with a single higher frequency neighbor. Similar results were reported by Grainger and Segui (1990) and Grainger (1990): Lexical decision latencies to words with higher frequency neighbors were delayed, but no cumulative inhibition was observed.

In contrast, Paap and Johansen (1994) reported not only an inhibitory neighborhood frequency effect in a lexical decision task, but also a cumulative inhibitory effect. In their study, lexical decision latencies to very high frequency, high-frequency, low-frequency, and very low frequency words were submitted to a regression analysis, which revealed that increases in the number of higher frequency neighbors were associated with increases in lexical decision latencies. However, the error rate for the very low frequency words was quite high (23.5%), which suggests that many of these words were either unknown or had uncertain spellings for their subjects. Because all of these very low frequency words had higher frequency neighbors, and because the mean response latency for these words was very slow (690 vs. 619 msec for the low-frequency words), it seemed likely that these words were the cause of both the inhibitory neighborhood frequency effect and the cumulative inhibitory effect. This suspicion was confirmed in a reanalysis of Paap and Johansen's data, in which the very low frequency words were excluded. In this analysis, both the basic effect and the cumulative effect disappeared.[1]

Sears et al. (1995) noted the apparent contradiction between the findings of Andrews (1989, 1992) and those of Grainger and colleagues. In particular, they noted that because low-frequency words with large neighborhoods are more likely to possess higher frequency neighbors than are low-frequency words with small neighborhoods, reports of facilitatory neighborhood size effects and inhibitory neighborhood frequency effects would seem to be contradictory. To shed more light on this issue, Sears et al. conducted six experiments in which neighborhood size and neighborhood frequency were factorially manipulated (i.e., words had either a small or a large neighborhood and had no neighbors of higher frequency or at least one higher frequency neighbor). The results of this investigation were straightforward: Facilitatory neighborhood size effects were consistently observed, but no inhibitory neighborhood frequency effects were obtained. In fact, Sears et al. repeatedly found that responses to words with higher frequency neighbors in lexical decision and naming tasks were actually faster than responses to words without higher frequency neighbors. In contrast to the conclusions of Grainger and colleagues, Sears et al. suggested that higher frequency neighbors do not inhibit lexical selection but instead, may actually facilitate it.

Grainger and Jacobs's (1996) Multiple Read-out Model

Recently, Grainger and Jacobs (1996) have described an activation-based model that can apparently accommodate not only facilitatory neighborhood size effects, but also inhibitory *and* facilitatory neighborhood frequency effects in lexical decision tasks. Grainger and Jacobs's *multiple read-out* model is based on the architecture of the interactive-activation model (McClelland & Rumelhart, 1981), in which a set of lexical and sublexical units accumulate activation over time. The distinguishing feature of this model is that it incorporates a number of variable decision criteria that can influence the nature of orthographic neighborhood effects in lexical decision. More specifically, the model possesses three decision criteria that influence the speed at which positive and negative lexical decisions are made. These are the M criterion, which is sensitive to single lexical unit activation; the Σ criterion, which is sensitive to global or summed lexical activation; and the T criterion, which is a temporal deadline used for negative responses. According to the model, if either the M criterion or the Σ criterion are reached before the T criterion is reached, a positive lexical decision response is made; otherwise, a negative response is made.

According to the model, when a word is presented, the lexical unit of the word and the lexical units of its neighbors are activated. A positive response can be made when the word's lexical unit reaches an activation threshold set by the M criterion (whereby lexical selection has been achieved), or when the total lexical activity generated by the word and its neighbors exceeds the current Σ criterion. The M criterion is fixed in the model, but the setting of the Σ criterion can be strategically adjusted, and its particular setting will determine whether positive responses are based more on lexical selection or on global lexical activity. More 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. Thus, the Σ criterion provides a means of making fast positive lexical decisions before a word is completely identified, under the assumption that lexical decision responses do not always require unique word identification (see, e.g., Seidenberg, Waters, Barnes, & Tannenhaus, 1984; Waters & Seidenberg, 1985).

According to Grainger and Jacobs (1996), the particular setting of the Σ criterion will determine whether inhibitory neighborhood frequency effects or facilitatory neighborhood size effects will be observed in a lexical decision experiment. More specifically, inhibitory neighborhood frequency effects will arise when the Σ criterion is set relatively high and "word" decisions are mainly based on lexical selection being completed (the M criterion). For example, according to the model, when the nonwords used in a lexical decision experiment are all very word-like (i.e., when they have large neighborhoods), inhibitory neighborhood frequency effects should occur, because subjects will be inclined to keep their Σ criterion high and, thus, the M criterion will tend to drive responding. That is, because nonwords with large neighborhoods will generate a great deal of lexical activity, it will be difficult to distinguish them from the words on the basis of this activity (i.e., via the Σ criterion). Thus, in order to avoid making errors, subjects will need to keep their Σ criterion high, meaning that most responses will be made on the basis of lexical selection (i.e., the M criterion being reached). Since the lexical selection process is strongly affected by inhibition from higher frequency neighbors, an inhibitory neighborhood frequency effect would result.

Conversely, when the nonwords used in an experiment are less word-like (i.e., when they have few neighbors), the Σ criterion will play more of a role, and a facilitatory neighborhood size effect should be observed. More specifically, because nonwords with few neighbors generate very little lexical activity, they are easy to distinguish from the words on the basis of global lexical activity. In these situations, subjects will be inclined to set their Σ criterion at a low level, because the lexical activity generated by a stimulus will reliably signal whether or not it is a word. Facilitatory neighborhood size effects will then occur, because words with large neighborhoods will generate more lexical activity than will words with small neighborhoods and, thus, words with large neighborhoods will cause the Σ criterion to be reached quite rapidly. These are essentially the patterns of effects that Grainger and Jacobs (1996) observed in their experiments, and thus, the model was successful in simulating these outcomes.

The strength of Grainger and Jacobs's (1996) model lies in its ability to account for reports of facilitatory neighborhood size effects (e.g., Andrews, 1989) and reports of inhibitory neighborhood frequency effects (e.g., Grainger et al., 1989) in lexical decision. Furthermore, Grainger and Jacobs have claimed that the model can even simulate the facilitatory neighborhood frequency effects reported by Sears et al. (1995). More specifically, when the model was tested on the identical set of words and nonwords used in one of Sears et al.'s experiments (Experiment 4A), settings of the Σ and T criteria could be found such that the model produced a small facilitatory neighborhood frequency effect in addition to a facilitatory neighborhood size effect.[2]

The Present Experiments

According to many models of word recognition, the lexical selection process is a core process in understanding words. The purpose of the present experiments was to evaluate the multiple read-out model's assumptions about this process—specifically, its assumptions about how the process is affected by orthographic neighborhood variables. To do so, we purposely selected tasks in which accurate responding could only be based on the successful completion of lexical selection (i.e., a single unit's activation level reaching the M criterion), and not on the amount of overall lexical activity (as, presumably, can occur, through use of the Σ criterion, in lexical decision tasks). Specifically, in Experiments 1 and 2, we used a perceptual identification task. In Experiment 3, we employed a semantic categorization task, in an attempt to extend the conclusions we drew from Experiments 1 and 2 and to reconcile those conclusions with Forster and Shen's (1996) results in their semantic categorization task.

The multiple read-out model makes two clear predictions in these types of tasks. Focusing initially on the perceptual identification task of Experiments 1 and 2, the first prediction is that facilitatory neighborhood size effects should not be observed. That is, accurate performance in the perceptual identification task requires that a single word be identified, and this will only be accomplished, according to the model, if the activation level in the appropriate lexical unit reaches the M criterion. In fact, the Σ and T criteria were omitted from the simulations of performance in Grainger and Jacobs's (1996) version of the perceptual identification task (i.e., their progressive demasking task). Because the Σ criterion is the component of the model that allows it to simulate facilitatory neighborhood size effects in lexical decision, the omission of this criterion eliminates the model's ability to produce similar effects in perceptual identification. Second, the model predicts that there will be clear inhibitory neighborhood frequency effects in a perceptual identification task. That is, according to the model, more lexical inhibition will arise for words with higher frequency neighbors than for words without higher frequency neighbors. As such, it would be less likely that the M criterion would be reached (and, hence, the word recognized) when processing words with higher frequency neighbors than when processing words without higher frequency neighbors.

Grainger and Jacobs (1996) were quite explicit with regard to both of these predictions. For example, when discussing the predictions of the model for a perceptual identification task (their Experiment 1A), they wrote:

> The model predicts the presence of an inhibitory effect of neighborhood frequency for the conditions tested in Experiment 1A. It also predicts the absence of a facilitatory effect of neighborhood density, because the latter is assumed to result from the use of the Σ and T criteria. By hypothesis, these criteria are not operational in the perceptual identification task. (p. 525)

Similar predictions would hold for the semantic categorization task (i.e., does the word name an animal?). This task also requires that words be uniquely identified, because accurate responding would depend on retrieving the appropriate meaning information. In the multiple read-out model, this can only be accomplished via the M criterion, and thus, as with the perceptual identification task, the Σ criterion should play no role, and facilitatory neighborhood size effects should not occur. However, intralexical competitive processes should produce an inhibitory neighborhood frequency effect.

We should note that other investigators have examined orthographic neighborhood effects on perceptual processing. The majority of these studies, however, have used the progressive demasking task rather than the standard perceptual identification task (e.g., Grainger & Jacobs, 1996; Grainger & Segui, 1990; Snodgrass & Mintzer, 1993). In this task, a trial consists of a continuous sequence of stimulus word and mask presentations. Over the course of a trial, the visibility of the stimulus word is gradually increased by decreasing the mask duration and increasing the word duration simultaneously. The subjects' task is to stop the demasking sequence as soon as they believe that they have identified the word. Response latencies are measured from the beginning of the sequence until the subjects' response (at which point they are requested to report the word).

Using this procedure, Grainger and Segui (1990) found that words with high-frequency neighbors took longer to identify than words with no higher frequency neighbors. Grainger and Jacobs (1996) have reported similar results. In terms of the multiple read-out model, these results would seem to confirm that the identification of words with higher frequency neighbors is relatively impaired. With respect to the effects of neighborhood size, however, the results, to date, have been less clear. Snodgrass and Mintzer (1993) found both facilitory and inhibitory neighborhood size effects and concluded that the specific nature of the neighborhood size effect was largely mediated by guessing strategies.

We did not choose to investigate orthographic neighborhood effects with a progressive demasking task, for two reasons. First, because there is currently very little data on the effects of orthographic neighbors in a standard perceptual identification task, we felt that it was neces-

Table 1
Mean Word Frequency (WF), Neighborhood Size (N), and Neighborhood Frequency (NBF) for the Stimuli Used in Experiments 1A and 1B

Neighborhood Frequency	Small WF	Small N	Small NBF	Large WF	Large N	Large NBF
	Low-Frequency Words					
No HF neighbors	19.6	2.0	7.0	21.8	7.5	11.5
HF neighbors	17.2	2.0	250.8	17.8	7.3	256.0
	High-Frequency Words					
No HF neighbors	100.4	2.1	30.0	104.9	7.4	41.3
HF neighbors	103.8	2.2	258.8	98.1	7.9	361.6

Note—HF, higher frequency. NBF refers to the mean frequency of the highest frequency neighbor.

sary to obtain converging evidence of the role of orthographic neighbors in data-limited tasks by using a data-limited task other than progressive demasking. Although Grainger and Jacobs (1996) have reported an inhibitory neighborhood frequency effect in a standard perceptual identification task, neighborhood frequency and neighborhood size were not factorially manipulated in that experiment (only small neighborhood words were used), and consequently the generality of their finding is somewhat limited.

Second, although all perceptual identification tasks are undoubtedly vulnerable to decision-making/guessing strategies, the perceptual demasking task seems to be especially susceptible because the task involves multiple exposures of the stimulus that, by necessity, add a time component to the trials. The result of this particular procedure would seem to be to encourage subjects to use information from the early exposures in order to generate and test hypotheses about the stimulus on the later exposures. Given that there will often be some ambiguity based on the early exposures, selection of the incorrect candidate for hypothesis testing may often occur. The result would be either an error (if the subject chose to respond with that candidate before the stimulus has been unambiguously identified) or a delay in producing the correct response (see Luo & Snodgrass, 1994, for an analogous argument applied to a picture identification task).

More important, selection of these candidates would probably be affected by neighborhood characteristics. For example, a word from a large neighborhood would seem to have a lower probability of being the first candidate selected for testing than would a word from a small neighborhood, just based on chance alone. As well, words having a higher frequency neighbor may have a lower chance of being selected, because the candidate selection process may be partly driven by familiarity. In both cases, the result would be the appearance of an inhibitory effect on lexical selection that was actually due to nonlexical factors. The standard perceptual identification task, of course, only involves a single presentation, and thus, although the task is undoubtedly not completely uncontaminated by guessing strategies, it should not be influenced by such hypothesis-testing strategies.

EXPERIMENT 1

In this experiment, we factorially manipulated word frequency, neighborhood size, and neighborhood frequency. The subjects attempted to identify briefly presented words that were forward and backward masked. The dependent variable was the percentage of correct identifications.

In addition, two visibility conditions were used. In Experiment 1A (the high-visibility condition), forward and backward mask durations of 500 msec were employed. In Experiment 1B (the low-visibility condition), the forward and backward mask durations were 42 msec. (Words were presented for 28 msec in both experiments.) On the basis of the results of a pilot experiment in which a number of mask durations were examined, we expected that performance would be approximately 80% in the high-visibility condition and approximately 50% in the low-visibility condition. Our purpose in manipulating stimulus visibility was to ensure that our test of the multiple read-out model's predictions (inhibitory neighborhood frequency effects and null neighborhood size effects), as well as our subsequent conclusions, were not limited to one particular level of identification performance.

Method

Subjects. Eighty-two undergraduate students from the University of Western Ontario participated in this experiment for course credit. Thirty-six participated in Experiment 1A (high-visibility condition), and 46 participated in Experiment 1B (low-visibility condition). All were native English speakers and reported that they had normal or corrected-to-normal vision.

Stimuli. All the stimuli were five-letter words. Descriptive statistics for these stimuli are listed in Table 1. (The complete set of experimental words used is presented in the Appendix.) Half of these words were high-frequency words (mean Kučera and Francis, 1967, normative frequency per million words of 101, range of 51–207), and the remainder were low-frequency words (mean normative frequency per million words of 19.1, range of 1–49).

The second factor manipulated was neighborhood size. Half of the words possessed a small neighborhood (i.e., at least one and no more than three neighbors); these had a mean neighborhood size of 2.1. The other half possessed a large neighborhood (i.e., at least five neighbors); these had a mean neighborhood size of 7.5. To be considered a neighbor of a target word, a word had to appear either in the Kučera 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 the words had at least one neighbor of higher frequency than themselves, whereas the other half of the words did not. The mean Kučera and Francis (1967) normative frequency of the highest frequency neighbor of each word was 310.2 for the high-frequency words with higher frequency neighbors and 253.4 for the low-frequency words with higher frequency neigh-

Table 2
Percentage of Correct Identifications in Experiment 1A

Neighborhood Frequency	Neighborhood Size Small	Neighborhood Size Large
Low-Frequency Words		
No HF neighbors	81.1	73.3
HF neighbors	76.4	75.1
High-Frequency Words		
No HF neighbors	79.6	78.1
HF neighbors	82.5	80.0

Note—HF, higher frequency.

bors. Finally, for both the low- and the high-frequency words with no higher frequency neighbors, the mean frequency of the highest frequency neighbor of each word was substantially lower than the mean frequency of the word itself.

Apparatus and Procedure. The stimuli were presented on a 17-in. color VGA monitor driven by a 80486-based microcomputer. The presentation of the stimuli was synchronized with the vertical retrace rate of the monitor (14 msec). At a viewing distance of 50 cm, the word stimuli subtended a visual angle of approximately 1.2°.

A trial sequence consisted of a forward mask, a word, and a backward mask. Each trial was initiated by a 1-sec 2000-Hz warning tone, after which a fixation point appeared at the center of the video monitor. Two seconds after the onset of the fixation point, a forward mask (#####) was presented directly above the fixation point for 500 msec in Experiment 1A (high-visibility condition) and for 42 msec in Experiment 1B (low-visibility condition). The forward mask was then removed, and a word was presented at the exact position of the mask (words were presented in uppercase letters because the descenders of lowercase letters are not masked by the # character). Words were presented for 28 msec (two vertical traces of the video monitor). Following the presentation of the word, a backward mask (#####) was presented for either 500 msec (Experiment 1A) or 42 msec (Experiment 1B). The backward mask was replaced by five question marks (?????), and 1 sec later, a prompt appeared at the bottom center of the video monitor ("What was the word?"). The subjects responded by typing in the word they had identified. They were instructed to use the editing keys on the keyboard to correct any typing errors, and to press the Enter key when they were satisfied with their response. There were no time constraints for responding.

Each subject completed 30 practice trials prior to the collection of data (these practice stimuli were not used in the experiment proper, and the data from these practice trials were not analyzed). During the practice trials, the duration of the word presentations was gradually decreased from 196 to 28 msec, to familiarize the subjects with the brief presentations. The 1st and 2nd stimuli were presented for 196 msec; the 3rd and 4th for 140 msec; the 5th and 6th for 84 msec; and the 7th, 8th, 9th, and 10th for 42 msec. The remaining 20 practice stimuli were presented for 28 msec, the same stimulus duration employed in the experiments. During the practice trials, the subjects were provided with feedback as to the accuracy of each response, and if an error was made, the correct word was presented on the computer monitor. No feedback was provided during the experimental trials.

Design. A 2 (word frequency: high, low) × 2 (neighborhood size: small, large) × 2 (neighborhood frequency: no higher frequency neighbors, higher frequency neighbors) factorial design was employed in both experiments. There were 15 words in each of the eight conditions, for a total of 120 words (see the Appendix). An additional 50 words of various word frequencies and neighborhood sizes were used as fillers. The order in which the 170 words were presented in the experiment was randomized individually for each subject.

Results

Tables 2 and 3 show the percentage of correct identifications for each of the eight stimulus conditions in Experiments 1A and 1B. The data for each experiment were submitted to a 2 (word frequency) × 2 (neighborhood size) × 2 (neighborhood frequency) repeated measures analysis of variance (ANOVA). Both subject (F_s) and item (F_i) analyses were carried out.[3]

High-visibility condition (Experiment 1A). Word frequency had a significant effect on identification performance [$F_s(1,35) = 10.75$, $MS_e = 85.13$, $p < .01$; $F_i(1,112) = 1.95$, $MS_e = 195.18$, $p = .16$]. Identification performance was superior for high-frequency words (80.0%), relative to low-frequency words (76.4%). The main effect of neighborhood size was also significant [$F_s(1,35) = 7.55$, $MS_e = 103.01$, $p < .01$; $F_i(1,112) = 1.66$, $MS_e = 195.18$, $p = .20$]. The overall neighborhood size effect was inhibitory: Words with large neighborhoods were identified less frequently (76.6%) than words with small neighborhoods (79.9%). The main effect of neighborhood frequency (the presence or absence of a higher frequency neighbor) was not significant ($F_s < 1$; $F_i < 1$). Thus, overall there was no inhibitory neighborhood frequency effect (78.5% for words with higher frequency neighbors and 78.0% for words with no higher frequency neighbors).

The interaction between word frequency and neighborhood size was not significant ($F_s < 1$, $F_i < 1$), since both the low- and the high-frequency words exhibited an inhibitory neighborhood size effect. The interaction between neighborhood size and neighborhood frequency also was not significant [$F_s(1,35) = 1.42$, $MS_e = 91.69$, $p > .20$; $F_i < 1$].

There was a marginally significant interaction between word frequency and neighborhood frequency [$F_s(1,35) = 3.39$, $MS_e = 76.56$, $p < .08$; $F_i < 1$]. The three-way interaction between word frequency, neighborhood size, and neighborhood frequency was also marginally significant [$F_s(1,35) = 4.06$, $MS_e = 63.86$, $p < .06$; $F_i < 1$]. An examination of Table 2 reveals the likely source of these marginal interactions. That is, there was an inhibitory neigh-

Table 3
Percentage of Correct Identifications in Experiment 1B

Neighborhood Frequency	Neighborhood Size Small	Neighborhood Size Large
Low-Frequency Words		
No HF neighbors	35.7	36.8
HF neighbors	42.4	44.4
High-Frequency Words		
No HF neighbors	42.3	47.5
HF neighbors	46.8	45.5

Note—HF, higher frequency.

borhood frequency effect for the low-frequency words with small neighborhoods (for the low-frequency words, the small-neighborhood words with higher frequency neighbors were identified less frequently than the small-neighborhood words with no higher frequency neighbors $[t(35) = 2.33, p < .05]$. However, this effect was confined to these low-frequency, small-neighborhood words. For the other word types, there was no hint of an inhibitory neighborhood frequency effect. In fact, the effects all went in the opposite direction, although none of these differences was statistically significant (all *p*s > .15).

Low-visibility condition (Experiment 1B). The main effect of word frequency was significant $[F_s(1,45) = 34.8$, $MS_e = 84.31, p < .001$; $F_i(1,112) = 2.86$, $MS_e = 335.13$, $p = .09]$. As was expected, identification performance was superior for high-frequency words (45.5%), relative to low-frequency words (39.8%). The main effect of neighborhood size was also significant $[F_s(1,45) = 4.07$, $MS_e = 68.38, p < .05$; $F_i < 1]$. However, in contrast to the results of Experiment 1A, the overall neighborhood size effect here was facilitatory, not inhibitory. That is, words with large neighborhoods were identified more frequently (43.5%) than words with small neighborhoods (41.8%). In addition, the main effect of neighborhood frequency was significant $[F_s(1,45) = 15.9, MS_e = 102.16$, $p < .001$; $F_i(1,112) = 1.58, MS_e = 335.13, p = .20]$. However, the neighborhood frequency effect was facilitatory, not inhibitory: Words with higher frequency neighbors were identified more frequently (44.7%) than words with no higher frequency neighbors (40.5%).

As was the case in Experiment 1A, the neighborhood size × neighborhood frequency interaction was not significant $[F_s(1,45) = 1.59, MS_e = 109.70, p > .20$; $F_i < 1]$, nor was the neighborhood size × word frequency interaction ($F_s < 1$; $F_i < 1$). The interaction between word frequency and neighborhood frequency was significant $[F_s(1,45) = 8.16, MS_e = 99.49, p < .01$; $F_i < 1]$ and was due to the fact that the facilitatory neighborhood frequency effect was much more pronounced for the low-frequency words. That is, for the low-frequency words, a facilitatory neighborhood frequency effect of 7.15% was observed $[F_s(1,45) = 34.2$, $MS_e = 69.12$, $p < .001$; $F_i(1,56) = 2.03$, $MS_e = 380.61$, $p = .16]$. For the high-frequency words, the neighborhood frequency effect was only 1.25% ($F_s < 1$; $F_i < 1$).

The three-way interaction between word frequency, neighborhood size, and neighborhood frequency was marginally significant $[F_s(1,45) = 3.16$, $MS_e = 103.36$, $p < .09$; $F_i < 1]$. The general trend in the data suggests that the low-frequency words exhibited a more consistent neighborhood frequency effect, and planned comparisons supported this interpretation. That is, for the low-frequency words, a facilitatory neighborhood frequency effect was evident for both the small- and the large-neighborhood words $[t(45) = 3.52, p < .01$; $t(45) = 4.06$, $p < .001$, respectively]. However, for the high-frequency words, only the small-neighborhood stimuli exhibited a facilitatory neighborhood frequency effect $[t(45) = 2.08$, $p < .05]$. For the large-neighborhood words, there was no facilitatory (or inhibitory) neighborhood frequency effect $[t(45) = 0.82$, n.s.]. What all of these interactions seem to suggest is simply that low-frequency words are more sensitive to neighborhood frequency manipulations than are high-frequency words, as has also been reported to be true in the naming and lexical decision tasks.

Combined analyses. A combined analysis of the data from Experiments 1A and 1B was performed to determine whether the neighborhood frequency effect and the neighborhood size effect significantly interacted with stimulus visibility (i.e., high visibility or low visibility, corresponding to Experiments 1A and 1B, respectively). In the subject analysis, the data were submitted to a 2 (visibility) × 2 (word frequency) × 2 (neighborhood size) × 2 (neighborhood frequency) mixed model ANOVA. The main effect of visibility was significant $[F_s(1,80) = 70.75, MS_e = 2{,}893.04, p < .001]$, as was the main effect of word frequency $[F_s(1,80) = 40.52$, $MS_e = 84.67, p < .001]$. The main effect of neighborhood frequency was significant $[F_s(1,80) = 8.58, MS_e = 104.52, p < .005]$ but the main effect of neighborhood size was not $[F_s(1,80) = 1.16, MS_e = 83.53, p > .25]$. More important, there were significant interactions between visibility and neighborhood frequency $[F_s(1,80) = 5.27, MS_e = 104.52, p < .05]$, and between visibility and neighborhood size $[F_s(1,80) = 12.21, MS_e = 83.53, p < .005]$. These interactions reinforce the observations made above; namely, (1) the neighborhood size effect changed from being inhibitory in Experiment 1A (the high-visibility condition) to being facilitatory in Experiment 1B (the low-visibility condition), and (2) although there was no significant effect of neighborhood frequency in the high-visibility condition, there was a significant facilitatory neighborhood frequency effect in the low-visibility condition.

The three-way interaction between visibility, word frequency, and neighborhood frequency was significant $[F_s(1,80) = 10.71, MS_e = 89.45, p < .005]$, which, as we have noted, reflected the fact that only the low-frequency words exhibited a significant facilitatory neighborhood frequency effect, and only in the low-visibility condition (Experiment 1B). Finally, the three-way interaction between word frequency, neighborhood size, and neighborhood frequency was also significant $[F_s(1,80) = 6.71$, $MS_e = 86.08, p < .05]$, owing to the varying nature of the neighborhood size and neighborhood frequency effects for the high-frequency and the low-frequency words. No other interactions were significant (all *p*s > .10). In the item analysis, the main effect of visibility was significant $[F_i(1,224) = 286.67, MS_e = 265.15, p < .001]$, as was the main effect of word frequency $[F_i(1,224) = 4.81$, $MS_e = 265.15$, $p < .05]$. No other effect was significant (all *p*s > .10).

Discussion

The important findings of Experiment 1 are as follows. First, we found that words with higher frequency neighbors did not consistently show lower levels of performance than did words without higher frequency neighbors (i.e., there was no evidence of an overall inhibitory neighborhood frequency effect). In Experiment 1A (high-visibility condition), only the low-frequency, small-neighborhood words showed any trend in this direction (with the other three word types actually showing trends in the opposite direction), and in Experiment 1B (low-visibility condition), words with higher frequency neighbors were identified more frequently than words with no higher frequency neighbors. It is worth noting that Grainger and Jacobs's (1996) inhibitory neighborhood frequency effect was based entirely on low-frequency words with small neighborhoods, because neighborhood frequency effects were examined only for those types of words. In any case, the present results clearly do not support the multiple read-out model's prediction that the existence of higher frequency neighbors globally inhibits perceptual identification.

Second, although the multiple read-out model did not predict any neighborhood size effects, the data from both experiments clearly show effects of neighborhood size on identification performance. More specifically, in the high-visibility condition, the neighborhood size effect was inhibitory (words with large neighborhoods were identified less frequently than words with small neighborhoods), and in the low-visibility condition, the neighborhood size effect was facilitatory (words with large neighborhoods were identified more frequently than words with small neighborhoods). The latter finding particularly raises problems for the multiple read-out model, because the mechanism used to explain facilitatory neighborhood size effects in lexical decision (the Σ criterion) is not operational in the perceptual identification task and, thus, the model would predict that there should be no effect of neighborhood size (and certainly not a facilitatory effect) in this task.

The change in the nature of the neighborhood size effect from an inhibitory one in the high-visibility condition to a facilitatory one in the low-visibility condition deserves additional comment. The inhibitory neighborhood size effect in the high-visibility condition is perhaps not too surprising, given that other investigators have reported inhibitory neighborhood size effects with the perceptual identification task (Snodgrass & Mintzer, 1993). However, like Snodgrass and Mintzer, we are reluctant to attribute this effect to lexical selection, because, on the basis of previous research, the lexical level influences of large neighborhoods appear to be facilitatory rather than inhibitory (Andrews, 1997). In this respect, the facilitatory neighborhood size effect (as well as the facilitatory neighborhood frequency effect) observed in the low-visibility condition nicely parallels the effects reported for English words in naming and lexical decision tasks (e.g., Andrews, 1989, 1992; Sears et al., 1995). Thus, we are inclined to believe that these previously reported phenomena reflect the "true" effects of these factors on the lexical selection process. That is, both being from a large neighborhood and having higher frequency neighbors seem to facilitate lexical selection.

This conclusion, of course, raises the question of why we did not observe a facilitatory neighborhood size effect in the high-visibility condition. A likely possibility is that the inhibitory neighborhood size effect was due to the effects of an "informed" guessing process that would be employed by the subjects when a word was generally largely visible. Our reasoning is as follows.

On the basis of the relatively high level of performance in the high-visibility condition, it appears that the stimuli were reasonably visible. Thus, on virtually all the trials, including the error trials, the subjects probably saw something. In particular, there may have been a reasonably large number of trials in which what they saw would not have allowed them to identify the target uniquely but would have allowed them to delimit a set of possibilities. Those possibilities would essentially be a subset of the words in the target's neighborhood. On those types of trials, one would expect that neighborhood characteristics would then become an important factor in determining the ultimate response. In particular, having a large neighborhood or having a higher frequency neighbor would disadvantage a word by making it less likely that the word itself would be given as the guessed response.

Support for this hypothesis can be gained by considering the nature of the error responses for words in Experiment 1A (high visibility) from large and small neighborhoods.[4] If this hypothesis is correct, the subjects should be producing error responses from the word's neighborhood (rather than correct guesses) much more often when the word has a large neighborhood than when the word has a small neighborhood. This was in fact the case. For the target words with large neighborhoods, 36.5% of the incorrect responses were neighbors of the target word, whereas for the target words with small neighborhoods, 20.7% of the incorrect responses were neighbors. (For the target words with higher frequency neighbors, 23.8% of the incorrect responses were higher frequency neighbors of the target word.) The likely implication is that there were, therefore, more accurate "guesses" for words from small neighborhoods than for words from large neighborhoods. If so, this would mean that the accuracy for small-neighborhood words was more inflated from informed guessing than the accuracy for large-neighborhood words was. Thus, any processing advantage that large neighborhoods would produce would be counteracted. (By the same argument, accuracy for words without higher frequency neighbors would also have been more inflated by guessing than would accuracy for words with higher frequency neighbors, making it even more surprising that we did not observe an inhibitory effect of neighborhood frequency in this experiment.)

In contrast, the low levels of stimulus visibility in the low-visibility condition may have created a situation in

Table 4
Mean Word Frequency (WF), Neighborhood Size (N), and Neighborhood Frequency (NBF1, NBF2) for the Stimuli Used in Experiments 2A and 2B

Stimulus Characteristic	Neighborhood Frequency		
	No HF Neighbors	One HF Neighbor	Two HF Neighbors
WF	43.5	44.8	42.2
N	8.0	8.0	8.2
NBF1	27.2	425.1	462.8
NBF2	12.0	43.4	219.2

Note—HF, higher frequency. NBF1 refers to the mean frequency of the highest frequency neighbor. NBF2 refers to the mean frequency of the second-highest frequency neighbor.

which there were substantially fewer trials in which the subjects had available to them a set of possibilities from which to make an informed guess (i.e., on most trials they either perceived the stimulus or saw very little, and thus they truly were just guessing). If so, there would have been relatively fewer trials on which the word's neighborhood's characteristics were affecting the decision-making/guessing process. Thus, the tendency for guesses based on neighborhood characteristics to counteract any lexical selection advantage that large neighborhoods or higher frequency neighbors provide would also have been somewhat less.

The data from the error trials of Experiment 1B (the low-visibility condition) are consistent with this hypothesis. For the target words with large neighborhoods, 17.9% of the incorrect responses were neighbors of the target word, and for the target words with small neighborhoods, 9.0% of the incorrect responses were neighbors. (For the target words with higher frequency neighbors, 10.8% of the incorrect responses were higher frequency neighbors of the target word.) Two points need to be made about these results. First, although the difference was smaller in the low-visibility condition than in the high-visibility condition, the proportion of incorrect responses that were neighbors was once again greater for large-neighborhood targets than for small-neighborhood targets. Thus, these results are further evidence that neighborhood characteristics do have the expected effect on guessing behavior. More important, in comparison with the high-visibility condition, these results represent substantial reductions in the percentage of errors that were neighborhood errors. As such, it follows that there were relatively fewer opportunities for informed guesses in the low-visibility condition than in the high-visibility condition. (Similarly, for the words with higher frequency neighbors, there was a reduction in the proportion of errors that were higher frequency neighbors of the target, from 23.8% in the high-visibility condition to 10.8% in the low-visibility condition.) Thus, whatever guessing advantage that words from small neighborhoods and words without higher frequency neighbors had would have been smaller in the low-visibility condition than in the high-visibility condition. The result, of course, would be that it would have been easier for any processing advantage that words from large neighborhoods or words with higher frequency neighbors had to emerge in Experiment 1B.

Regardless of whether this explanation of the difference in the neighborhood size effects between Experiments 1A and 1B is correct, one rather important fact remains. We have obtained absolutely no evidence for an inhibitory neighborhood frequency effect in either of these experiments. Although the absence of such an effect in Experiment 1A might not raise problems for the multiple read-out model, the presence of a facilitatory neighborhood frequency effect in Experiment 1B is exactly the opposite of what the model predicts. As such, we felt it was important to turn our full attention to this effect and to ask whether we could replicate these findings. Accordingly, in Experiment 2, neighborhood frequency effects were again examined in a perceptual identification task, and to further increase the generalizability of our replication, a completely different set of stimuli was used (i.e., four-letter words). By demonstrating that a facilitatory neighborhood frequency effect also occurs for four-letter words, the effect witnessed in Experiment 1B cannot be attributed to some peculiarity of the set of five-letter words used in that experiment.

EXPERIMENT 2

In this experiment, the focus was solely on the effects of neighborhood frequency in perceptual identification. Three stimulus conditions were created: words that possessed (1) no higher frequency neighbors, (2) one higher frequency neighbor, or (3) two higher frequency neighbors. Word frequency and neighborhood size were equated across these conditions. In addition, two visibility conditions were used. In Experiment 2A (the high-visibility condition), forward and backward mask durations of 500 msec were employed, which were identical to the mask durations used in Experiment 1A. In Experiment 2b (the low-visibility condition), the forward and backward mask durations were 42 msec, identical to the mask durations used in Experiment 1B. Our purpose in

Table 5
Percentage of Correct Identifications in Experiments 2A (High Visibility) and 2B (Low Visibility)

	Neighborhood Frequency		
Visibility	No HF Neighbors	One HF Neighbor	Two HF Neighbors
High visibility	73.1	73.7	76.1
Low visibility	26.8	37.7	33.1

Note—HF, higher frequency.

manipulating stimulus visibility was to determine whether we would replicate the pattern of results observed in Experiments 1A and 1B: no neighborhood frequency effect under the high-visibility condition and a facilitatory neighborhood frequency effect under the low-visibility condition.

Method

Subjects. Sixty undergraduate students from the University of Western Ontario participated in this experiment for course credit. Twenty-eight participated in Experiment 2A (high-visibility condition), and 32 participated in Experiment 2B (low-visibility condition). All were native English speakers and reported that they had normal or corrected-to-normal vision. None of these individuals had participated in Experiment 1.

Stimuli. The stimuli were low- and high-frequency four-letter words, with a mean Kučera and Francis (1967) normative frequency per million words of 43.5 (range of 4–97). The descriptive statistics for these stimuli are listed in Table 4. (The complete set of words used in Experiments 2A and 2B is presented in the Appendix.) All the words had large neighborhoods, each word possessing at least five neighbors, with a mean neighborhood size of 8.0. None of these words had been used in the previous experiments.

The single factor manipulated in this experiment was neighborhood frequency—that is, the presence or absence of a higher frequency neighbor in the word's orthographic neighborhood. There were three neighborhood frequency conditions. Target words possessed (1) no higher frequency neighbors, (2) one higher frequency neighbor, or (3) two higher frequency neighbors. For the words with no higher frequency neighbors, the mean Kučera and Francis (1967) frequency of the highest frequency neighbor of each word was substantially lower than the mean target frequency. For words with one higher frequency neighbor, the mean Kučera and Francis normative frequency of the highest frequency neighbor of each word was substantially higher than the mean target frequency. The words in this condition only possessed one neighbor substantially higher in frequency, although the next highest frequency word in the neighborhood was often very similar in frequency to the target. Finally, for words with two higher frequency neighbors, the mean Kučera and Francis normative frequency of the two highest frequency neighbors of each word was substantially higher than the mean target frequency.

Apparatus and Procedure. The apparatus and procedure were identical to those of Experiment 1. Each subject completed 30 practice trials prior to the collection of data. The procedure for these practice trials was identical to that of Experiment 1.

Design. The three neighborhood frequency conditions (no higher frequency neighbors, one higher frequency neighbor, two higher frequency neighbors) produced a one-factor repeated measures design. There were 25 words in each of the three conditions, for a total of 75 words. The order in which the 75 words were presented in the experiment was randomized individually for each subject.

Results

Table 5 shows the percentage of correct identifications for each of the three stimulus conditions of Experiments 2A and 2B. The data were submitted to a one-factor ANOVA. Again, both subject (F_s) and item (F_i) analyses were carried out.

High-visibility condition (Experiment 2A). Overall identification performance was 74.3%, which was very similar to the performance observed in the large neighborhood conditions of Experiment 1A (76.6%). The neighborhood frequency manipulation did not have a significant effect on identification performance [$F_s(2,54) = 1.42$, $MS_e = 50.11$, $p > .20$; $F_i < 1$]. That is, as was the case in Experiment 1A, no inhibitory (or facilitatory) neighborhood frequency effect was observed in the high-visibility condition.

On error trials, 25.9% of the incorrect responses were neighbors of the target word. For the words with higher frequency neighbors (i.e., for the words in the one higher frequency neighbor condition or the two higher frequency neighbors condition), 18.8% of the incorrect responses were higher frequency neighbors of the target word.

Low-visibility condition (Experiment 2B). In the low-visibility condition, overall identification performance dropped to 32.5%, comparable with the performance observed in the analogous conditions of Experiment 1B (40.6%). As in Experiment 1B, there was a significant effect of neighborhood frequency [$F_s(2,62) = 13.50$, $MS_e = 70.59$, $p < .001$; $F_i(2,72) = 2.05$, $MS_e = 364.02$, $p = .13$]. As is shown in Table 5, words with higher frequency neighbors were identified correctly more frequently than were words with no higher frequency neighbors, replicating the facilitatory neighborhood frequency effect observed in Experiment 1B.

Planned comparisons confirmed that identification performance was superior for both sets of words with higher frequency neighbors. That is, words with one higher frequency neighbor were identified more frequently than words with no higher frequency neighbors [$t(31) = 5.16$, $p < .001$], and words with two higher frequency neighbors were identified more frequently than words with no higher frequency neighbors [$t(31) = 3.73$, $p < .01$]. There was a marginally significant difference in identification performance between the one higher frequency neighbor and the two higher frequency neighbor conditions [$t(31) = 1.88$, $p = .07$].

On error trials, 14.3% of the incorrect responses were neighbors of the target word. For the words with higher frequency neighbors, 9.8% of the incorrect responses were higher frequency neighbors of the target word. Two points need to be made about these error results. First, as compared with the high-visibility condition (Experiment 2A), the percentage of errors that were neighborhood errors was reduced (25.9% vs. 14.3%), which again suggests that there were relatively fewer opportunities for informed guesses in the low-visibility condition. Second, for the words with higher frequency neighbors, there was a reduction in the proportion of errors that were higher frequency neighbors of the target, from 18.8% in the high-visibility condition to 9.8% in the low-visibility condition. As a consequence, any processing advantage that words with higher frequency neighbors have would have been more likely to emerge in the low-visibility condition, because it would have been less likely to be counteracted by informed guessing based on neighborhood characteristics.

Combined analyses. A combined analysis of the data from the low-visibility and the high-visibility conditions was performed to determine whether the neighborhood frequency effect significantly interacted with stimulus visibility. The subject means were submitted to a 2 (visibility) × 3 (neighborhood frequency) mixed model ANOVA. The main effect of visibility was significant [$F_s(1,58) = 78.5$, $MS_e = 994.68$, $p < .001$], as was the main effect of neighborhood frequency [$F_s(2,116) = 9.03$, $MS_e = 61.05, p < .001$]. More important, the interaction between visibility and neighborhood frequency was significant [$F_s(2,116) = 6.79, MS_e = 61.05, p < .01$], which reflected the fact that a facilitatory neighborhood frequency effect only occurred in the low-visibility condition. In the item analysis, the main effect of visibility was significant [$F_i(1,144) = 212.53$, $MS_e = 307.56, p < .001$], but the main effect of neighborhood frequency and the interaction between visibility and neighborhood frequency were not [$F_i(2,144) = 1.50$, $MS_e = 307.56$, $p > .20$; $F_i(2,144) = 1.13$, $MS_e = 307.56, p > .30$, respectively].

Finally, in order to address any lingering concerns about the generalizability of the facilitatory neighborhood frequency effect across items in the low-visibility condition, an item analysis of the combined data from the low-frequency, large-neighborhood conditions of Experiment 1B and the analogous conditions of Experiment 2B (no higher frequency neighbors, one higher frequency neighbor) was performed. The main effect of experiment was marginally significant [$F_i(1,76) = 3.73$, $MS_e = 349.89, p = .057$], since the items used in Experiment 1B were identified more frequently (40.6%) than the items used in Experiment 2B (32.2%). The main effect of neighborhood frequency was also significant [$F_i(1,76) = 4.61$, $MS_e = 349.89, p < .05$], since words with higher frequency neighbors were identified more frequently than words with no higher frequency neighbors (41.0% and 31.8%, respectively). Equally important, the interaction between experiment and neighborhood frequency was not significant ($F_i < 1$). An identical item analysis of the combined data from Experiments 1A and 2A (the high-visibility conditions) revealed no effect of experiment or neighborhood frequency and no interaction (all $F_i < 1$).

Discussion

The results of this experiment are quite clear: In the high-visibility condition, no facilitatory or inhibitory neighborhood frequency effects were observed, but in the low-visibility condition, it was easier to identify words with higher frequency neighbors than words with no higher frequency neighbors. Apart from replicating the pattern of neighborhood frequency effects observed in Experiments 1A and 1B with a completely different set of stimuli, these findings clearly demonstrate that inhibitory neighborhood frequency effects are not readily obtained in the standard perceptual identification task.

EXPERIMENT 3

As has been noted, according to the multiple read-out model, facilitatory neighborhood size effects in lexical decision are not due to the lexical selection process per se but are instead due to a variable response criterion that is sensitive to the degree of overall lexical activation (the Σ criterion). According to the model, words with large neighborhoods will generate more lexical activity than will words with small neighborhoods and will, therefore, allow this lowered Σ criterion to be reached more rapidly. This will in turn produce a facilitatory neighborhood size effect. However, in tasks such as perceptual identification, where a word must be uniquely identified, the Σ criterion is assumed to play no role, because the degree of overall lexical activation does not provide any information as to the correct response. In these situations, the lexical selection processes must run to completion, with the intralevel inhibition between the lexical units in the model strongly influencing performance. Thus, the model makes two predictions in this task: (1) There will be no neighborhood size effect, and (2) because higher frequency neighbors of a target word will inhibit the lexical unit of the target word, there will be an inhibitory neighborhood frequency effect.

Clearly, our perceptual identification data do not provide any support for these predictions. However, as has been noted, the perceptual identification task is not the only task in which the multiple read-out model makes these predictions. In fact, it makes these same predictions for any task in which a word must be uniquely identified. For example, a semantic categorization task (e.g., does the word name an animal?) would also require that words must be uniquely identified, because accurate responding would depend on retrieving the appropriate meaning information. Thus, as with the perceptual identification task, the Σ criterion should play no role in a semantic categorization task, and facilitatory neighborhood size ef-

Table 6
Mean Word Frequency (WF), Neighborhood Size (N), and Neighborhood Frequency (NBF) for the Stimuli Used in Experiment 3

Neighborhood Frequency	Small WF	Small N	Small NBF	Large WF	Large N	Large NBF
No HF neighbors	23.7	1.9	6.0	23.4	7.3	13.2
HF neighbors	21.7	2.2	309.8	20.1	7.2	277.5

Note—HF, higher frequency. NBF refers to the mean frequency of the highest frequency neighbor.

fects should not occur. Instead, intralexical competitive processes should produce an inhibitory neighborhood frequency effect.

Forster and Shen (1996) have recently examined neighborhood size and neighborhood frequency effects in a semantic categorization task. They reported that neither neighborhood size nor neighborhood frequency seemed to affect performance, owing to the fact that neither effect was statistically significant in their item analyses. These results led Forster and Shen to suggest that neighborhood size effects (and, presumably, neighborhood frequency effects) in lexical decision tasks may be due to decision biases, and not to the lexical selection process. It should be noted, however, that Forster and Shen observed both a significant facilitatory neighborhood size effect and a significant facilitatory neighborhood frequency effect in their subject analyses. Furthermore, this significant neighborhood size effect was obtained despite the fact that their manipulation of neighborhood size was quite weak—the "large-neighborhood" words in their experiments had only three or four neighbors. (In contrast, the large-neighborhood words in the present experiments had at least five neighbors.) Thus, although Forster and Shen's results strongly suggest that facilitatory neighborhood size effects and facilitatory neighborhood frequency effects are at least partially caused by decision biases in the lexical decision task, it would seem premature to conclude that neither factor (neighborhood size, neighborhood frequency) actually affects lexical selection.

In Experiment 3, we reexamined the question of whether there are neighborhood size and neighborhood frequency effects in a semantic categorization task. Our purpose in conducting this experiment was twofold. First, as was argued above, the multiple read-out model clearly does predict an inhibitory neighborhood frequency effect in a task of this sort, and we wished to provide another evaluation of that model's predictions. Second, because Forster and Shen's (1996) conclusions were based on a null effect of a rather weak manipulation in a negatively biased analysis (i.e., the statistical power of item analyses is reduced, owing to a greatly deflated alpha value when items have not been selected randomly; Cohen, 1976), we felt it was important to revisit this issue. Accordingly, in this experiment, neighborhood size and neighborhood frequency were factorially manipulated (i.e., words had either a small or a large neighborhood and had no neighbors of higher frequency or at least one higher frequency neighbor), and the task was to make an *animal* or a *non-animal* judgment. As in Forster and Shen's experiment, only a single category was used throughout the present experiment, and we considered only the response latencies to nonexemplars (*no* decisions). The reason for this is that latencies to animal exemplars are contaminated by both semantic priming effects and category typicality effects, whereas response latencies to nonexemplars are not. These two modifications simplify the interpretation of the task latencies (Bradley & Forster, 1987).

Method

Subjects. Thirty-five undergraduate unpaid student volunteers from the University of Calgary participated in this experiment. All of the subjects were native English speakers and reported that they had normal or corrected-to-normal vision. None of these individuals had participated in the previous experiments.

Stimuli. The experimental words (the nonexemplars) were all of low frequency, with a mean Kučera and Francis (1967) normative frequency of 22.2 per million words (range of 1–49). The descriptive characteristics of these stimuli are listed in Table 6. (The complete set of experimental words used in Experiment 3 is presented in the Appendix.) To create a suitably large set of experimental words, we used both four- and five-letter words, since we could not create a large set of animal names that were only four, or only five, letters in length. The majority of the experimental words (69%) had been used in Experiments 1 and 2.

Words with large neighborhoods had at least five neighbors, with a mean neighborhood size of 7.2. Words with small neighborhoods had at least one neighbor and no more than three neighbors, with a mean neighborhood size of 2.0. The mean Kučera and Francis (1967) normative frequency of the highest frequency neighbor of each word was 293.6 for the words with higher frequency neighbors and 9.6 (substantially lower than the mean target frequency) for the words with no higher frequency neighbors.

There were 24 words in each of the four experimental conditions, and an additional 96 four- and five-letter words that were animal names. Thus, a total of 192 words were presented in the experiment. The animal names included mammals, fish, reptiles, birds, amphibians, and insects, but excluded humans. This point was made clear in the instructions to the subjects (e.g., the subjects were told that the word *ant* should be categorized as an *animal*).

Apparatus and Procedure. The stimuli were presented on a 17-in. color VGA monitor driven by a Pentium-class microcomputer. The subjects indicated the semantic category of the stimuli (animal or non-animal) by pressing one of two buttons on a response box. The subjects used their preferred hand to respond to the *animal* items. The presentation of the stimuli was synchronized with the vertical

Table 7
Mean Semantic Categorization Latencies (in Milliseconds) and Error Rates (in %) in Experiment 3

Neighborhood Frequency	Small *M*	Small %	Large *M*	Large %
No HF neighbors	622	5.7	607	4.0
HF neighbors	624	3.7	599	4.3

Note—HF, higher frequency.

retrace rate of the monitor (14 msec), and response latencies were measured to the nearest millisecond. At a viewing distance of 50 cm, the word stimuli subtended a visual angle of approximately 1.2°.

Each trial was initiated by a 1-sec 2000-Hz warning tone, after which a fixation point appeared at the center of the video monitor. One second later, the stimulus was presented directly above the fixation point in uppercase letters. The subject's response terminated the stimulus display. The next trial was initiated after a timed interval of 2 sec.

Each subject completed 20 practice trials prior to the collection of data (these practice stimuli were not used in the experiment proper). The practice trials consisted of 10 animal words and 10 non-animal words. The order in which the 192 stimuli were presented in the experiment was randomized individually for each subject. The subjects were provided with a rest period after every 48 trials.

Design. A 2 (neighborhood size: small, large) × 2 (neighborhood frequency: no higher frequency neighbors, higher frequency neighbors) factorial design was employed. Since only the negative trials (i.e., the non-animals) were considered in the analysis, these factors were only relevant to the negative stimuli.

Results

Table 7 shows the mean response latencies for correct responses and the mean error rates in each of the four conditions of this experiment. As has been noted, as in the Forster and Shen (1996) study, only the data from the negative trials were analyzed. Response latencies were submitted to both subject (F_s) and item (F_i) analyses. Only the latencies for correct responses were analyzed.

In the analysis of response latencies, the main effect of neighborhood size was significant in both the subject [$F_s(1,34) = 26.19$, $MS_e = 501.59$, $p < .001$] and the item [$F_i(1,92) = 9.53$, $MS_e = 979.77$, $p < .01$] analyses. Words with large neighborhoods were responded to an average of 20 msec faster than words with small neighborhoods. Thus, these data indicate that a facilitatory neighborhood size effect (in particular, an effect that is significant even in an item analysis) can be observed in a semantic categorization task. The main effect of neighborhood frequency was not significant ($F_s < 1$; $F_i < 1$); nor was the interaction between neighborhood size and neighborhood frequency [$F_s(1,34) = 1.49$, $MS_e = 675.32$, $p > .20$; $F_i < 1$]. Thus, overall there was, once again, no inhibitory neighborhood frequency effect (and, in fact, the trend was in the opposite direction).

Error rates were also submitted to a 2 (neighborhood size) × 2 (neighborhood frequency) repeated measures ANOVA. The main effect of neighborhood size was not significant ($F_s < 1$); nor were the main effect of neighborhood frequency [$F_s(1,34) = 2.31$, $MS_e = 12.07$, $p > .10$], and the interaction between neighborhood size and neighborhood frequency [$F_s(1,34) = 2.86$, $MS_e = 15.66$, $p > .10$].

Discussion

The first point to make about these results is that, once again, no inhibitory neighborhood frequency effect was observed, despite the fact that, as in the previous experiments, the multiple read-out model predicts such an effect. Indeed, for the words with large neighborhoods, the trend that was observed was, once again, a trend for a facilitatory neighborhood frequency effect.

The second point to make about these results is that there was a facilitatory neighborhood size effect, which was significant in both the subject and the item analyses. Recall that Forster and Shen (1996) found significant facilitatory neighborhood size effects in their subject analyses, but not in their item analyses, which led them to suggest that the neighborhood size effect may be partially or entirely due to decision biases. However, given the results of this experiment, we suspect that the main reason Forster and Shen did not obtain a significant facilitatory neighborhood size effect in their item analyses was because their neighborhood size manipulation was relatively weak.

More specifically, in Forster and Shen's (1996) experiments, four neighborhood size conditions were used: Words possessed no neighbors, one neighbor, two neighbors, or three or four neighbors. In contrast, in this experiment, a larger range of neighborhood sizes was employed: Words with large neighborhoods had a mean neighborhood size of 7.2, and words with small neighborhoods had a mean neighborhood size of 2.0. In fact, relative to other studies in which neighborhood size has also been manipulated as a dichotomous variable (large vs. small), Forster and Shen's stimuli would all be classified as small-neighborhood words. In Andrews's (1989) Experiment 1, for example, the words with small neighborhoods had a mean neighborhood size of 2.3, whereas the words with large neighborhoods had a mean neighborhood size of 13.0. Similarly, in Sears et al.'s (1995) Experiment 1, words with small and large neighborhoods had mean neighborhood sizes of 3.6 and 10.5, respectively. Thus, using the criteria employed by other investigators, there were no large-neighborhood words in the Forster and Shen experiments, and consequently the statistical power to detect a neighborhood size effect would necessarily have been somewhat small. When that fact is coupled with the negative bias inherent in item ANOVAs (Cohen, 1976), it is not surprising that Forster and Shen were unable to obtain a significant neighborhood size effect in their item analysis.

It should also be noted again that the appearance of a facilitatory neighborhood size effect in this task also is contrary to the predictions of the multiple read-out model. The mechanism hypothesized to produce facilitatory neighborhood size effects in the model (the Σ criterion) is assumed to play no role in a semantic categorization task. Overall then, the results of this experiment, like the results of the previous experiments, would appear to cause considerable problems for the multiple read-out model.

GENERAL DISCUSSION

In terms of the initial motivation for this research, this set of experiments contains a number of important findings. First, we have found no evidence whatsoever of an inhibitory neighborhood frequency effect on perceptual identification performance. In Experiment 1A, the identification of words with higher frequency neighbors was equivalent to that of words with no higher frequency neighbors, and in Experiment 1B, the neighborhood frequency effect was facilitatory, not inhibitory. This pattern of results was replicated in Experiment 2, using a different set of stimuli. These results are difficult to reconcile with the multiple read-out model, because the model predicts that any effect of neighborhood frequency in the perceptual identification task should be inhibitory. That is, although the absence of such an effect might not raise problems for the model, there is no simple way in which the presence of a facilitatory neighborhood frequency effect could be accommodated.

Second, the results of Experiment 1B indicate that identification performance was facilitated when a word possessed a large neighborhood, a result that is also inconsistent with the multiple read-out model. More specifically, because the mechanism used to explain facilitatory neighborhood size effects in lexical decision (the Σ criterion) is not operational in the perceptual identification task, the model would predict that there should be no effect of neighborhood size (and certainly not a facilitatory effect) in this task.

Third, we found that words with large neighborhoods are responded to more rapidly in the semantic categorization task. Although this result would seem to be inconsistent with that of Forster and Shen (1996), as was noted earlier, the inconsistency is more apparent than real. That is, these investigators did find a significant facilitatory neighborhood size effect in their subject analyses in both of their experiments. More important, Forster and Shen's neighborhood size manipulation was more limited than ours, since only words with what are typically thought of as small neighborhoods were employed in their experiments. Thus, the source of the discrepancy between our results and those of Forster and Shen is probably due to this difference in the neighborhood size manipulation.

In contrast, the facilitatory neighborhood size effect observed in Experiment 3 is clearly inconsistent with the multiple read-out model. According to the model, facilitatory neighborhood size effects should not occur in a semantic categorization task, because the mechanism used to explain such effects (the Σ criterion) is not operational in tasks of that sort.

Although our data do not completely rule out the multiple read-out model, they do pose a fairly serious challenge for it. The multiple read-out model specifically predicts that facilitatory neighborhood size effects and facilitatory neighborhood frequency effects will only be found in lexical decision tasks. Thus, our demonstrations of such effects in perceptual identification and semantic categorization tasks directly contradict the model's predictions.

Recall that one of the strengths of the multiple read-out model is its ability to accommodate facilitatory neighborhood size effects and both inhibitory and facilitatory neighborhood frequency effects in lexical decision tasks. According to the model, facilitatory neighborhood size effects are not actually lexical selection effects but are instead due to a variable response criterion that is sensitive to the overall degree of lexical activation (the Σ criterion). Under the right circumstances, the placement of this criterion could even produce a facilitatory neighborhood frequency effect. The inhibitory neighborhood frequency effect, in contrast, is assumed to result from intralevel competitive processes that occur during the process of lexical selection, and thus this phenomenon is assumed to be a "true" lexical selection effect. Our results challenge this interpretation, and in so doing, they question the assumption that the neighborhood size effect is not a "true" lexical selection effect.

With respect to the issue of the locus of the neighborhood size effect, it should first be noted that facilitatory neighborhood size effects have considerable generalizability (see Andrews, 1997, for a review). They have now been observed in lexical decision, naming, perceptual identification (Experiment 1B), and semantic categorization (Experiment 3) tasks. The fact that the neighborhood size effect persists in tasks that require semantic access reinforces the notion that it reflects basic lexical selection processes. Facilitatory neighborhood size effects have also been observed when target words are embedded in sentences and eye fixations are measured. That is, Lima and Inhoff (1985) found that first fixation durations and gaze durations for words with large neighborhoods (*low-constraint* words in their study) were shorter than those for words with small neighborhoods (*high-constraint* words). This entire pattern of results would seem to argue that the advantage enjoyed by words with large neighborhoods is due to the nature of our mental architecture, rather than being an artifact of laboratory procedures (i.e., the decision-making operations in a lexical decision task).

In contrast, inhibitory neighborhood frequency effects have not been consistently observed in these same tasks (Andrews, 1997). For example, the majority of studies that have used the naming task have reported facilitatory neighborhood frequency effects, the single exception being Carreiras, Perea, and Grainger (1997). In this study, the neighborhood frequency effect was inhibitory for words with small neighborhoods but slightly facilitatory for words with large neighborhoods, producing an overall null effect. Furthermore, although the original reports of such effects in lexical decision were quite clear (e.g., Grainger et al., 1989), as has been noted, in subsequent studies null and even facilitatory neighborhood frequency effects have been reported (e.g., Forster & Shen, 1996; Sears et al., 1995). Inhibitory neighborhood frequency effects have also been elusive in semantic categorization tasks. In Carreiras et al.'s semantic categorization experiment, there was an inhibitory neighborhood frequency effect, but only for words with large neighborhoods. For words with small neighborhoods, a significant facilitatory neighborhood frequency effect was observed (which, as was noted, was also the case in Forster and Shen's study). Finally, in the present experiments, no inhibitory neighborhood frequency effects were observed in either the semantic categorization or the perceptual identification task, despite the fact that, according to the multiple read-out model, it is in these types of tasks that the effect should be most pronounced.

Clearly, the literature to date offers no compelling support for two of the basic premises of the multiple read-out model—that facilitatory neighborhood size effects are due to a variable response criterion instead of lexical selection processes, and that the lexical selection processes themselves produce an inhibitory neighborhood frequency effect. Consequently, it seems appropriate to also consider the present findings in terms of other models of word recognition.

As previously noted, facilitatory neighborhood size effects are, in general, incompatible with serial-search models of lexical selection (i.e., Forster, 1976; Paap et al., 1982). Facilitatory neighborhood frequency effects are clearly incompatible with such models as well, because higher frequency neighbors encountered during the serial search would delay (not facilitate) lexical selection. Consequently, the results of our perceptual identification experiments would be difficult to explain in terms of serial-search models.

Given the difficulties that facilitatory neighborhood effects pose for serial-search models, most investigators have turned to activation-based models of word recognition to explain these effects. In particular, Andrews (1989) has argued that facilitatory neighborhood size effects can be accommodated by the interactive-activation model (McClelland & Rumelhart, 1981). In the interactive-activation model, the presentation of a word activates the lexical units of the word and its neighbors, and lexical selection is achieved when the word's lexical unit reaches a critical activation threshold. For high-frequency words, which have high resting activation levels, lexical selection can often occur through direct bottom-up activation alone. Thus, the activation of neighbors will play very little role in the process.

On the other hand, because low-frequency words have lower resting activation levels, the build-up of bottom-up activation in a word's lexical unit will be much slower. In these situations, the activation of the word's lexical unit can be facilitated by the reciprocal activation mechanism embodied in the model. That is, the partially activated units corresponding to the word and its neighbors send excitatory feedback back down to their sublexical units, which, in turn, send activation back up to the lexical units, increasing the activation of the lexical units themselves. These lexical-to-sublexical activation reverberations continue until the activation of the word's lexical unit exceeds the activation threshold, at which point lexical selection is achieved. The point to note, then, is that low-frequency words with large neighborhoods would benefit more from this reciprocal activation process than would low-frequency words with small neighborhoods, because a greater number of lexical units would participate in the reciprocal activation process. Thus, the expectation would be that words with large neighborhoods would be recognized faster than words with small neighborhoods, as Andrews (1989) and others have observed. As a result, words with large neighborhoods should be reported more readily in perceptual identification tasks and responded to more rapidly in semantic categorization tasks, because the lexical selection process, the process that is being facilitated, is intrinsic to those tasks.

Similarly, the facilitatory neighborhood frequency effects witnessed in Experiments 1B and 2B could also be explained by the same mechanism. That is, higher frequency neighbors, which possess high resting levels of activation, could produce stronger top-down activation, which could accelerate the reciprocal activation process and facilitate the processing of words with higher frequency neighbors (Sears et al., 1995). As such, all the results of the present experiments would seem to be compatible with the interactive-activation model.

In fact, it would appear that the interactive-activation model could accommodate either facilitatory or inhibitory orthographic neighborhood effects, depending on whether top-down excitatory feedback or lexical inhibition is assumed to govern the process. The idea would simply be that, if the model were to explain the numerous reports of facilitatory orthographic neighborhood effects, the role of lexical inhibition in the model would have to be presumed to be minor in comparison with the role of lexical–sublexical feedback activation. The extent to which these modifications would harm the model's ability to account for other word recognition phenomena would, of course, need to be evaluated.

What should also be noted is that, because the multiple read-out model is an offshoot of the interactive-activation model, the multiple read-out model itself could also account for the present data if some of its assumptions were changed. In the case of the multiple read-out model, however, these changes would clearly not be minor ones. The model is based specifically on the assumptions that (1) intralevel inhibition is a major component of lexical processing and (2) lexical–sublexical feedback activation is not. With these assumptions, the model was able to account for the inhibitory neighborhood frequency effects that Grainger and colleagues reported. Furthermore, these assumptions then made it necessary to account for all facilitatory effects in terms of the activity of external factors, such as the Σ criterion in a lexical decision task. As such, to change either of these assumptions would be to change the model itself. Thus, at present, it makes sense to distinguish between the two models in terms of their ability to account for the present results. The interactive-activation model appears to be able to do so quite adequately. The multiple read-out model does not.

Finally, as the reader will note, the models discussed here are all based on the principle that it is the existence of large neighborhoods or higher frequency words that causes the mental representations to be established in particular ways, which then leads to the observed effects. That is, the assumption is that it is the existence of certain sets of words in the language that causes our mental representations to be set up in a particular way. What is worth keeping in mind is that the factors we manipulated here (neighborhood size and neighborhood frequency) are correlational. Thus, the direction of causation of neighborhood effects may actually be somewhat different than that proposed by all these models. That is, it may be the case, for example, that words with large neighborhoods may be easier to process, not because their large neighborhood is somehow mentally represented, but rather because they have an orthographic structure that is "user friendly" and, hence, many words may have been created to exploit that structure. If so, there would seem to be a whole class of models, based not on lexical representations but on orthographic representations and orthographic processing, that may actually provide the best account of neighborhood effects.

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.

BRADLEY, D. A., & FORSTER, K. I. (1987). A reader's view of listening. *Cognition*, **25**, 103-134.

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-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, **12**, 335-359.

COHEN, J. (1976). Random means random. *Journal of Verbal Learning & 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). New York: Academic Press.

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

FORSTER, K. I., & SHEN, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, & 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., & SEGUI, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, **47**, 191-198.

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 & Verbal Behavior*, **15**, 263-265.

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

LIMA, S. D., & INHOFF, A. W. (1985). Lexical access during eye fixations in reading: Effects of word-initial letter sequence. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 272-285.

LUO, C. R., & SNODGRASS, J. G. (1994). Competitive activation model of perceptual interference in picture and word identification. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 50-60.

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

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

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

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.

SEIDENBERG, M. S., WATERS, G. S., BARNES, M. A., & TANNENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, **23**, 383-404.

SMITH, J. E. K. (1976). The assuming-will-make-it-so fallacy. *Journal of Verbal Learning & Verbal Behavior*, **15**, 262-263.

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

WATERS, G. S., & SEIDENBERG, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition*, **13**, 557-572.

WIKE, E. L., & CHURCH, J. D. (1976). Comments on Clark's "The Language-as-Fixed-Effect Fallacy." *Journal of Verbal Learning & Verbal Behavior*, **15**, 249-255.

NOTES

1. We are grateful to Ken Paap for making these data available to us.

2. According to Grainger and Jacobs (1996), the reason that the model can account for facilitatory neighborhood frequency effects is because words with higher frequency neighbors produce more lexical activity than do words with no higher frequency neighbors, even when the neighborhood sizes have been equated. Thus, if the subjects in the Sears et al. (1995) experiments used the Σ criterion for responding, facilitatory neighborhood frequency effects, as well as facilitatory neighborhood size effects, could be observed. A closer look at Grainger and Jacobs's (1996) simulations, however, suggests that Grainger and Jacobs's claim might be a bit strong. In their simulations of Experiment 4 from Sears et al., for example, Grainger and Jacobs reported the results of simulations of all the nonword data, but of only a subset of the word data. As a second example, consider their reported simulation of the results of Sears et al.'s Experiment 5, in which nonwords with large neigh-

borhoods were used. As noted, according to the multiple read-out model, when the nonwords are all very word-like (i.e., have large neighborhoods), inhibitory neighborhood frequency effects (and null neighborhood size effects) would be expected, because subjects would be inclined to keep their Σ criterion high, meaning that their M criterion would drive responding. However, in this experiment, a facilitatory neighborhood size effect was observed, as was a trend toward a facilitatory neighborhood frequency effect. Nonetheless, Grainger and Jacobs found that they could simulate these results by adjusting the Σ and T criteria. What Grainger and Jacobs did not report, however, were the results of the simulations for the nonwords in Experiment 5. Because nonword latencies and error rates are strongly affected by modifications of the Σ and T criteria and because the specific purpose of Experiment 5 was to induce changes in these criteria by introducing difficult nonwords, the question of the adequacy of the simulation of nonword performance would seem to be a key one. Thus, at present, it is somewhat difficult to accurately evaluate the model's ability to simulate the full pattern of data from Sears et al.'s experiments.

3 As it turns out, no effects were significant in item analyses in either of the perceptual identification experiments reported in the present paper (Experiments 1 and 2). We do not, however, regard this as an issue. The reason is that although Clark (1973) has argued that items, as well as subjects, should be considered as a random factor in these types of analyses, it is seldom the case that the selection of items is ever random in any sense of the term. That is, typically, the items used in these types of experiments have been selected because they satisfied an extensive set of criteria, which is certainly the case here (see, e.g., Table 1). Consequently, as Wike and Church (1976) and others (Cohen, 1976; Keppel, 1976; Smith, 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 because of a greatly deflated alpha value). Lingering concerns that readers might have about the generalizability of our perceptual identification results across items are addressed by the fact that the neighborhood frequency effects in Experiment 2, using an entirely new set of items, are quite similar to those in Experiment 1. Moreover, the neighborhood frequency effect was significant in an item analysis when the stimuli from Experiments 1B and 2B were combined.

4. We thank Leann Stadtlander for suggesting this analysis.

APPENDIX

Items Used in Experiments 1A and 1B

High Frequency/Small Neighborhood/No Higher Frequency Neighbors: ALIVE, ANGLE, AWARE, CIVIL, COAST, CROWD, DREAM, DRINK, FLOOR, FRESH, METAL, NORTH, STAFF, TRADE, VALUE

High Frequency/Small Neighborhood/Higher Frequency Neighbors: ALONE, BEGIN, CLEAN, DEPTH, HEART, IDEAL, MOUTH, PEACE, ROMAN, SPEND, THICK, UNITY, WOMEN, WORTH, YOUTH

High Frequency/Large Neighborhood/No Higher Frequency Neighbors: BREAK, BROWN, CARRY, CLASS, DRAWN, MODEL, REACH, RIVER, SCALE, SHARE, STOCK, SWEET, TASTE, TRAIN, WATCH

High Frequency/Large Neighborhood/Higher Frequency Neighbors: BEACH, FIGHT, GLASS, LOWER, OLDER, PLANE, PROVE, ROUND, SCORE, SHAPE, SHORE, SIGHT, SOUND, STAGE, STORE

Low Frequency/Small Neighborhood/No Higher Frequency Neighbors: ANGER, BLAST, BOOST, EXACT, FLEET, GLOOM, HARSH, LABEL, LODGE, MERCY, PANEL, PLEAD, SAUCE, SOLAR, SPRAY

Low Frequency/Small Neighborhood/Higher Frequency Neighbors: AWAKE, COUNT, DENSE, FEAST, FLOUR, LOYAL, MANOR, MARSH, MAYOR, REACT, TOKEN, TREAT, VOCAL, WEAVE, YIELD

Low Frequency/Large Neighborhood/No Higher Frequency Neighbors: BAKER, BORED, EAGER, FREED, GRACE, JOLLY, LUNCH, METER, PITCH, PORCH, SCOUT, SHINE, SILLY, SNAKE, WIPED

Low Frequency/Large Neighborhood/Higher Frequency Neighbors: BAKED, BLANK, GRADE, PAINT, PEACH, PLATE, POKER, PRIME, SHOCK, SLACK, SPICE, SPIKE, SPILL, TIGHT, TRACE

Items Used in Experiments 2A and 2B

No Higher Frequency Neighbors: BOAT, BOWL, DIED, DIRT, DUKE, FLAT, FLOW, GIFT, JOKE, JUMP, LOAN, MILK, PATH, PLOT, PUSH, RAFT, ROCK, SHIP, SKIN, SLAB, SOAP, SOIL, SPAN, TEAM, TUBE

One Higher Frequency Neighbor: CALM, CAMP, COAT, CORN, DISH, DOWN, FAIR, FIST, FOOT, HAZE, HERO, HORN, KING, LINK, LOOP, LUNG, MASK, RAIN, RULE, SAFE, TOOL, WARM, WASH, WEAK, YARD

Two Higher Frequency Neighbors: BURN, DUST, EASE, FAST, FOOL, FORT, HARM, HEAT, HERD, JEEP, LOAD, MAID, MEEK, NOSE, PAIN, PAIR, SHOE, SHOP, SNOW, SOLD, TOAD, TOUR, VAST, WEEP, ZONE

APPENDIX (Continued)

Items Used in Experiment 3

Small Neighborhood/No Higher Frequency Neighbors: ATOM, BLAST, BOMB, BOOST, CRIB, DEBT, EXACT, FLEET, FOLK, GLAD, GLOOM, IDLE, INCH, IRON, LABEL, LODGE, MERCY, MYTH, NAVY, PANEL, SAUCE, SOLAR, SPRAY, ZINC

Small Neighborhood/Higher Frequency Neighbors: AWAKE, AXLE, COUNT, DENSE, FEAST, FLOUR, FOAM, HAUL, HELM, HOLY, KNEE, KNIT, KNOT, LOYAL, MANOR, MAYOR, POEM, REACT, SHUT, TOKEN, TREAT, VARY, VERB, YIELD

Large Neighborhood/No Higher Frequency Neighbors: BAKER, BORED, BOWL, COPY, DIRT, DRILL, FREED, GIFT, GRACE, JOKE, JOLLY, JUMP, LUNCH, METER, PITCH, PLOT, PUSH, RAFT, SHINE, SILLY, SLAB, SOAP, TUBE, WIPED

Large Neighborhood/Higher Frequency Neighbors: BAKED, BLANK, DISH, GLUE, GRADE, HARM, KISS, LOAD, LOFT, LOOP, PAINT, PLATE, POKER, PRAY, SHOCK, SLACK, SPICE, SPILL, TIGHT, TOOL, TRACE, WEAK, YARD, ZONE

(Manuscript received January 17, 1997;
revision accepted for publication August 18, 1998.)