Do postonset segments define
a lexical neighborhood?

ROCHELLE S. NEWMAN
University of Maryland, College Park, Maryland

and

JAMES R. SAWUSCH and PAUL A. LUCE
State University of New York, Buffalo, New York

Previous research has demonstrated that the number and frequency of lexical neighbors affects the perception of individual sounds within a nonword in a phoneme identification task. In the present research, the issue of what items should be considered part of a word's neighborhood was explored. These experiments, in which both lexical decision and phoneme identification tasks were used, demonstrate that lexical neighborhood effects are not limited to words that match the target item syllable initially (the cohort). Words that differ from a target only in their first phoneme influence the process of lexical access. This argues against the notion that word onsets serve a unique or special purpose in word recognition.

A number of studies have demonstrated neighborhood effects in spoken word recognition (Goldinger, Luce, & Pisoni, 1989; Luce, Pisoni, & Goldinger, 1990). The neighborhood effect refers to the finding that perception of a word (or a nonsense word) is affected by the number and frequency of occurrence of similar-sounding words in the language. For example, Luce and Pisoni (1998) found that words that are similar to many words in the language (e.g., *cat*, *kit*, *pat*, and *cap*) are responded to more slowly than words with fewer neighbors, in both lexical decision and naming tasks. Goldinger et al. (1989) demonstrated similar effects in a priming task, and Cluff and Luce (1990) found that identification performance for spondees (two-syllable words composed of two monosyllabic words, such as *baseball*) was influenced by the size of the neighborhood for each syllable. Finally, Vitevitch and Luce (1999) demonstrated neighborhood effects in *same–different* matching and in semantic categorization (animate vs. inanimate), showing that these effects occur even in tasks that do not focus attention on the sound pattern of the words.

Newman, Sawusch, and Luce (1997) demonstrated that neighborhoods also play a role in phonetic perception.

Their study was modeled after experiments by Ganong (1980), in which the lexical effect was examined. The lexical effect refers to the finding that perception of an ambiguous phonetic segment is affected by the lexical status of the syllable in which that segment occurs. Ganong presented listeners with pairs of voice onset time (VOT) continua, such that only one endpoint in each continuum constituted a real word in English. For example, one pair of continua included items ranging from the word *beef* to the nonword *peef* and from the nonword *beace* to the word *peace*. Listeners were more likely to label the ambiguous items in the middle of the two series as a real word. Thus, an item ambiguous between *b* and *p* would be labeled as a *b* in the *beef* to *peef* series, but as a *p* in the *beace* to *peace* series. Newman et al. (1997) demonstrated a similar effect of lexical neighborhood. They used pairs of series in which none of the four endpoints were real words but one endpoint of each pair was similar to a greater number of real words in the language. For example, one pair of series included items ranging from *gice* to *kice* and *gipe* to *kipe*, where *gice* is similar to more words than is *kice* and *kipe* is similar to more words than is *gipe*. Listeners classified ambiguous items as being whichever endpoint made them similar to more actual words in English.

The results in Newman et al. (1997) clearly demonstrate that lexical neighborhoods can influence phoneme perception. However, the precise nature of how the lexicon exerts this influence is still unclear. Most important, it is not entirely clear what constitutes the set of words that influence phoneme perception. Newman et al. (1997) followed the one-phoneme change rule described by Luce (1987), in which all words that differ from the target word by the addition, deletion, or substitution of a single

phoneme were considered to be neighbors of the target. Although it would be preferable to base the neighborhood computation on an independent set of data that established the similarity among words, there are problems with using existing data sets. For example, similarity scaling judgments of consonants and vowels in CVCs might be expected to reflect both the details of the particular talker's voice and the similarity of the phonetic segments. Consequently, Newman et al. (1997) argued that it was more appropriate to use the coarse approximation of a one-phoneme change rule that is talker independent than to use detailed similarity scaling data that might be talker specific. With the one-phoneme change rule, *scat*, *at*, *cash*, and *sat* are all neighbors for *cat*. Perhaps more controversially, *cap*, *cot*, and *sat* are all treated equivalently as neighbors, despite the differences in the location of the phonetic mismatch.

An alternative approach suggests that only words that match the target in the initial phoneme will influence further processing of the target. That is, *cap* would be activated by the initial /k/ and the vowel of *cat*. Neither *sat* nor *scat* would be activated, because the initial /s/ in both does not match the acoustic-phonetic information for the initial /k/ in the target. This view was proposed as part of the original cohort work of Marslen-Wilson and Welsh (1978) but has received support from a wide array of studies (see Marslen-Wilson, 1990, for a discussion). For example, Marslen-Wilson and Zwitserlood (1989) reported that rhymes have a much smaller priming effect than do whole words, suggesting that word onsets are particularly salient. Although more recent versions of the cohort theory have tempered its claims somewhat, placing more emphasis on the degree of overlap between an input and a lexical representation (Marslen-Wilson, 1987), they continue to put special emphasis on the initial portion of a word, so that an item that mismatches at word onset would likely be only weakly activated, if at all. One exception is that when an item mismatches a lexical entry by only a single feature and results in a nonword (rather than another word), it may still prime the target word (Marslen-Wilson, Moss, & van Halen, 1996).

The importance of word onsets has also been reported in the domains of mispronunciation detection and false memories. Studies have reported that both children and adults are more likely to detect mispronunciations that occur early in words than those occurring late, suggesting that more attention is paid to word onsets than to later syllables (Cole, Jakimik, & Cooper, 1978; Walley, 1987; but see Cole & Perfetti, 1980; van Donselaar, 1996). Similarly, Westbury, Buchanan, and Brown (2002) reported more false alarms in a memory task for CVC words that matched studied items in their initial CV than for words that matched in their final VC. They suggested that lexical "activation spreads in a serial manner, with more emphasis on the beginning of the word than on the end" (p. 637).

In contrast to these results, Connine, Blasko, and Titone (1993) found comparable amounts of priming for nonwords that differed from a target word in their initial phoneme and those that differed in their medial phoneme. These findings suggest that initial phonemes receive no greater weighting than do later phonemes in lexical activation and imply that neighborhoods should be based on global similarity, regardless of the positions in which mismatches occur.

There may be gradations between these two extremes. One example, in addition to the revised versions of the cohort model described above, is the neighborhood activation model (Luce & Pisoni, 1998). In this model, information in the signal activates lexical entries, starting with the information underlying the initial phoneme of the target. Since speech information occurs over time, the information in the initial segment can have a greater influence on word recognition than does the information that occurs later, because it starts the process of activating a neighborhood and words that become activated sooner will compete more effectively than words that are only activated later. However, a word that does not match the initial phoneme of the target can still be activated as part of the neighborhood and compete for recognition. That is, as the segment information for *cat* is accumulated, it will match both *scat* and *sat* sufficiently to partially activate them. The words *scat* and *sat* are part of the lexical neighborhood for the target *cat* but will receive activation only later in processing than will the neighbor *cap*. This approach would suggest that words that match at word onset have a greater initial influence on target word perception than do words that mismatch at onset but that the latter words should still influence perception. In line with this approach, Allopenna, Magnuson, and Tanenhaus (1998) used an eye-tracking methodology to examine which potential competitors received activation during lexical processing. They found earlier activation for items that matched a target word at onset but found that other items began to be activated soon thereafter and remained active longer.

In the present work, we attempt to contrast these different approaches to word recognition and to examine more deeply the issue of which words are part of the lexical neighborhood that is activated in speech perception. We will use the term *neighborhood* in two ways. The first is a generic usage referring to the words in the mental lexicon that are similar to a target and are partially activated during the process of word recognition. The second usage involves a specification of the similarity among words. Here, a lexical neighborhood consists of all of the words that differ from the target in a single phoneme, regardless of where the one phoneme difference occurs. This is one of the formal definitions of neighborhood described by Luce (1987). The two uses of the term *neighborhood* should be clear from the context in which they are used. The term *cohort* will be used to refer to words that match the target at word onset and differ from the target in one phoneme. The difference between a neighborhood and a cohort revolves around the influence of words that differ from the target in their initial phoneme but match it in all other respects. Our use of the term co-

hort here is more restrictive than Marslen-Wilson and Welsh's (1978), who used it to refer to all of the words in the mental lexicon that begin with the same phoneme as the target. However, our use does preserve a key distinction between neighborhoods and cohorts, which is whether partial activation of lexical candidates requires that the candidate words match the target at onset. Consequently, we will use the term *cohort* as we have defined it, for clarity of presentation.

In Experiments 1 and 2, we used a lexical decision task to compare performance for targets that vary in cohort size (items that match at word onset) with performance for targets that vary in neighborhood size. In Experiments 3 and 4, we searched for neighborhood effects in the task used by Newman et al. (1997), using items with no cohort. That is, these targets have different-sized neighborhoods, but all the neighbors mismatch the targets on the first phoneme. If perception of these target items is still influenced by the lexical neighborhood, it would be strong evidence that at least some words that differ in initial position are activated and affect lexical processing. To anticipate our results, the data across these experiments suggest that, indeed, all words similar to a target word become activated during the process of word recognition: We consistently find effects of neighbors that differ from the target word at word onset.

EXPERIMENT 1

In the first experiment, we addressed the issue of which items should be considered part of a lexical neighborhood, using a lexical decision task. Early versions of cohort theory (e.g., Marslen-Wilson & Welsh, 1978) proposed that in word recognition, the initial phoneme would lead to activation of a cohort of lexical candidates. As additional phonemes were processed, initially active but now inconsistent word candidates would be deactivated until only one word remained. According to this type of proposal, only words that match the target carrier in the first phoneme should influence listeners' responses. The alternative, embodied in models such as NAM (Luce & Pisoni, 1998), is that neighborhood activation is based on the overall similarity between the lexical entries and the target. Thus, *rice* is part of the activated neighborhood for the nonword *gice*, even though they do not share the initial phoneme. This approach still allows for the possibility that the neighbors that mismatch at onset will play a lesser role in processing than will the neighbors that match at onset. This possibility has not been explored in most experimental manipulations of lexical neighborhood.

In the present experiment, we used a lexical decision task with two sets of nonword syllables as targets. Neighborhood properties were calculated for these sets in two different ways. One method of calculation included all the items that differed from a target item by one phoneme in any position: a one-phoneme substitution, addition, or deletion. We refer to this as an *overall neighbor-hood analysis*. Our second method of calculation involved cohorts and included those items that differed in one phoneme from the target and matched in their initial phonemes. Thus, this second method of calculation ignores any contribution from that subset of similar words that differ in their initial phoneme. Since these items already differ from the target word at the point when lexical activation begins, there may be no reason to activate them. This might not be the case in fluent speech, where the onset of a word might be ambiguous. However, in the lexical decision task, words are presented in isolation, and the word onset is clear.

Nonwords were selected such that one set of items had higher neighborhood values if only items matching word initially were treated as members of the neighborhood. We will refer to this set as the *onset-match set* to emphasize that the neighborhood computation required that the words in the neighborhood match the initial phoneme of the target. The other set of items had higher neighborhood values when all the neighbors (as defined previously) were included and will be referred to as the *full-neighborhood set*. In the lexical decision task, any factor that makes a nonword more like a word should increase the errors that listeners make to the nonwords, increase the amount of time listeners take to respond to the nonwords, or both. This implies that the set of nonwords responded to more slowly will indicate which method of neighborhood computation is more accurate. If the cohort (those items that match word initially) have the only influence on perception or have a greater influence on perception than do words that do not match initially, the onset-match set of nonsense syllables is more wordlike and should produce slower response times (RTs) or more errors than does the full-neighborhood set. Conversely, if all the words that differ by only one phoneme from the target are part of the neighborhood and are activated to a nearly equivalent extent, the items in the full-neighborhood set would be more wordlike and should be responded to more slowly and/or less accurately than those in the onset-match set.

Method

Participants. Thirty-one members of the University of Iowa community participated in this experiment for course credit. All the participants were right-handed and native speakers of English, with no reported history of a speech or hearing disorder. The data from an additional 3 participants were not included in the analysis, because these individuals were left-handed.

Stimuli. Two groups of 20 CVC nonsense words were selected as stimuli. These items are listed in Appendix A. The two sets were matched for positional phoneme probabilities. That is, each phoneme that occurs syllable initially in one set occurs syllable initially the same number of times in the other set. This same matching was done for the vowel and the final consonant. However, matching on biphones (sequences of two phonemes) was not possible while maintaining other constraints. The average neighborhood values of both sets, via both methods of calculation, are shown in Table 1.

Neighborhood calculations were performed in the same manner as that in Newman et al. (1997). Each nonword was compared with real words in an on-line dictionary. For the full-neighborhood set, we defined an item's neighbors as being every real word that dif-

Table 1
Average and Total Frequency-Weighted Neighborhood Density
for the Two Sets of 20 Nonwords Used in Experiment 1

Set	Onset-Matching Neighborhood		Full Neighborhood	
	Average	Total	Average	Total
Onset match	9.3	186.3	11.9	237.5
Full neighborhood	4.2	83.6	17.0	339.3

Note—*Average* is the average neighborhood density per word; *total* sums across the 20 words. The columns on the left include only those neighbors that match at word onset; the columns on the right include all neighbors.

fered from the target word by only a single phoneme substitution, addition, or deletion. For the onset-match set, we defined an item's neighbors as being every real word that differed from the target word by only a single phoneme substitution, addition, or deletion and also matched the target word in its first phoneme. After determining the number of neighbors for each item by each approach, neighbors were weighted by their log-transformed frequencies[1] and were summed to yield a frequency-weighted neighborhood density. This weighting ensures that words that are more common in the language contribute more to a lexical neighborhood than do rare words. It is worth noting, however, that Marslen-Wilson (1990) has found evidence that the effect of the frequency of neighbors may play a role on word recognition only at relatively early stages of processing; lexical decision tasks in his study did not show effects of the frequency of the neighbors (although the number of neighbors in that study was not controlled). Only those neighbors with a familiarity index of 6.0 or greater (Nusbaum, Pisoni, & Davis, 1984) were included, to avoid basing our calculations on unfamiliar words that were unlikely to be in our listeners' lexicons. This method of calculating frequency-weighted neighborhood density has been described previously by Luce (1987). This resulted in two sets of words. If only those words that match the target item in their first phoneme are considered neighbors, the onset-match set had a higher frequency-weighted neighborhood density. Conversely, if all the words that differ from the target in a single phoneme are included, the items in the full-neighborhood set are more wordlike, as shown by their higher frequency-weighted neighborhood density.

An additional 40 real words were selected as foils. These items were all neighbors of the nonsense items. Approximately one third of the words differed from a nonsense syllable in their first consonant, one third in their vowel, and one third in their final consonant.

All items were recorded by a female native speaker of English (R.S.N.), using a Shure SM81 microphone and a Mackie 1202 VLZ mixer/amplifier. They were recorded at a 44.1-kHz sampling rate, with 16-bit quantization, and were stored on computer disk.

Procedure. The participants were asked to decide whether each item was a real word in English or a nonsense item and to respond by pressing the appropriate button on a computer-controlled response box. Both speed and accuracy were emphasized. Since we were particularly interested in the nonword items, the right button was used for nonword responses, and the left button was used for word responses.

Presentation pace depended on the listeners' response speed. The next trial began 1 sec after the listener had responded or after an interval of 3 sec from stimulus onset had elapsed, whichever came first. The listeners' responses and RTs were recorded for each stimulus. Any trial on which the RT was more than three standard deviations from a participants' mean RT was removed from analysis. These exceptionally long RTs probably reflect momentary inattention to the task. They are also very unlikely to reflect the on-line use of the information in the stimulus to identify words. This resulted in the removal of 36 trials across the 31 participants, or 1.7% of the data.

A Macintosh 7100/AV computer controlled stimulus presentation and response collection. The stimuli, which were stored on disk, were presented binaurally through AudioTechnica ATH-M40fs headphones. Prior to the experiment, the listeners received a practice block of 10 items (half words, half nonwords). They then heard each of the 80 target block items once, with the order randomized.

The items in this study were not selected randomly; they were very closely matched on phonotactic probability, as well as being specifically chosen so as to have different values by the two methods of calculating neighborhoods. For this reason, the appropriate analysis is by participants alone; item analyses are not appropriate (Raaijmakers, 2003; Raaijmakers, Schrijnemakers, & Gremmen, 1999). However, since we realize that many readers would prefer to see both participant and item analyses, we will include them both for completeness but will focus on the results from the participant analysis. For all significant analyses, η^2 values were computed to indicate the magnitude of the effect. As a correlational-type indicator of *proportion of variance accounted for*, an η^2 value of .01 indicated a small effect, a value of .06 represented a medium effect, and a value of .14 or greater represented a large effect size (Cohen, 1988).

Results

Although we expected the items that were similar to real words to be difficult, we wanted to ensure that all the items were generally heard correctly. For that reason, all analyses were performed twice. In the first analysis, all the items were compared on both accuracy and RTs. In the second analysis, any item that received fewer than 80% correct responses across participants was removed, and only the remaining items were compared. The pattern of results was identical across these two types of analysis; those reported below include all the items.

There was a significant difference between the two sets in terms of accuracy, so that the items in the full-neighborhood set were less accurate than those in the onset-match set [$t(30) = 2.98, p < .01$]. This accuracy difference was significant by participants, but not by items [$t(38) = 0.52, p > .10$]. The effect size was large ($\eta^2 = .228$). The effect is shown in Figure 1. The participants had a harder time deciding that the nonsense items were not words when they were higher in their overall neighborhood value than when they were similar to more words that matched their onset. This suggests that all words similar to a target word become activated during the course of recognition, regardless of whether they match word initially. There was no significant effect of the type of neighborhood computation in the RT data [$t(30) = -0.57, p > .10$ by participants; $t(38) = -0.51, p > .10$ by items].

These data must be interpreted with caution. One concern is that these effects appeared in the accuracy data, rather than in the RT data. Second, the lower accuracy for the full-neighborhood set was found only in the participant analysis. Most previous experiments in which neighborhood effects have been examined with lexical decision have found effects in listeners' RTs, whereas this experiment found a difference in accuracy. One possibility is that this finding may be an indication of the listeners' strategy. Although the instructions asked the listeners to respond as quickly as possible, we also emphasized that they should make as few errors as possible.

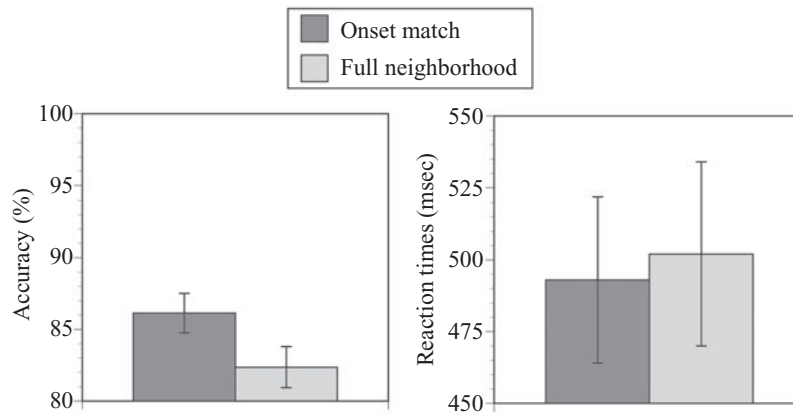Accuracy and reaction time data, Experiment 1

Figure 1. Lexical decision accuracy and reaction times for nonwords with different types of neighborhoods from Experiment 1. Error bars reflect standard errors.

Within the context of these instructions, it is plausible that the listeners would try to respond rapidly and, thus, make errors on difficult trials.

It is possible that the overall pattern of results shows an influence of both the full neighborhood and the cohort of the onset-matched items. If all items (those that match at onset and those that do not) are activated during word recognition, the set of words with higher neighborhood values would be responded to more slowly. This is what was found in the participant analysis. It is also plausible that although all neighbors were activated, those that matched in onset were activated earlier and may have had a slightly stronger influence than those that did not match initially. This would tend to diminish the overall difference between the two sets of items. This explanation still emphasizes that all the members of the neighborhood, including those that do not match the target at onset, are activated. However, it remains the case that the results of Experiment 1 offer, at best, partial support for the neighborhood approach and that they do not show whether the cohort does or does not influence the activation of lexical candidates (beyond representing lexical candidates that are activated first). To address this, in Experiment 2, we examined these two factors separately from one another.

EXPERIMENT 2

In Experiment 2, the lexical decision task was used with new sets of nonword syllables. One group of listeners heard two sets of nonword items that were matched in terms of their overall neighborhoods. The two sets differed in terms of neighborhood values when only items matching at word onset were considered. We will refer to this as the *onset-match* condition, using the same terminology as in Experiment 1. This set differs from that in Experiment 1

in that the two sets differed only in onset-matched neighborhoods. Their full neighborhoods were the same. A second group of listeners responded to two sets of nonwords that differed in their full neighborhoods but were matched in their cohorts. Again, following the terminology of Experiment 1, we refer to this as the *full-neighborhood* condition. In contrast to Experiment 1, the two sets of nonwords were matched with respect to neighborhoods that matched the target in their first phoneme. Thus, across these two groups of listeners, the onset-matched and the full-neighborhood computations were manipulated independently.

If, as the results from Experiment 1 suggest, all items that differ from a target word by a single phoneme contribute approximately equally to a lexical neighborhood, we would expect to find a significant effect of neighborhood for the full-neighborhood group. In particular, those items that are high in terms of their overall neighborhood values should be responded to more slowly (or less accurately) than those items that are low in overall neighborhood values. If only those neighbors that match at word onset contribute to lexical neighborhood, we would expect to find a significant effect of neighborhood for the onset-match group, so that those items that have more neighbors that match at onset should be responded to more slowly (or less accurately) than those items that have few such neighbors.

We might also find that both approaches are correct to some degree. That is, if onset-match neighbors have a greater effect than do neighbors that mismatch at target word onset, but the latter still contribute to lexical neighborhood to some degree, we might expect to find significant effects for both groups of listeners. The relative strength of these effects would be an indication of the degree to which these two types of neighbors differ in their influence on perception.

Method

Participants. Seventy-four members of the University of Maryland and University of Buffalo communities participated in this experiment for course credit or a cash payment. All the participants were right-handed and native speakers of English, with no reported history of a speech or hearing disorder. The data from an additional participant were excluded from the analysis because the listener was given incorrect instructions (a reversal of the hand-to-button mapping). Forty-one of these listeners were assigned to the onset-match condition, and 33 to the full-neighborhood condition. The data from 1 participant in the full-neighborhood condition were excluded because that participant failed to reach a preset accuracy criterion (this participant performed at a below-chance level for the nonword items, whereas all the other listeners had accuracy scores above 75%).

Stimuli. Four sets of CVC nonsense syllables were selected as targets. These items are listed in Appendix B. For the onset-match condition, two sets of 28 nonwords were selected. These items matched in terms of overall neighborhood values and in terms of positional phoneme probabilities (as in Experiment 1). For the full-neighborhood condition, two sets of 25 nonwords were selected. These items matched in terms of the number and frequency of neighbors with the same onset as the target and in terms of positional phoneme probabilities. Given this close matching across sets in both conditions, item analyses are not appropriate (Raaijmakers, 2003; Raaijmakers et al., 1999); we include them for completeness, as in Experiment 1, but focus on the more appropriate participant analyses. The average neighborhood values of all four sets of nonwords, using both methods of calculation, are shown in Table 2.

Neighborhood calculations were performed in the same manner as in Experiment 1. An additional 56 real words were selected as foils for the onset-match condition, and 50 words were chosen as foils for the full-neighborhood condition. All the items were recorded by the same speaker and in the same manner as in Experiment 1.

Procedure. The procedure was identical to that in Experiment 1. The participants were asked to decide whether each item was a real word in English or a nonsense word and to respond by pressing the appropriate button on a computer-controlled response box. Both speed and accuracy were emphasized. Any trial on which the RT was more than three standard deviations from a participant's mean RT was removed from analysis. This resulted in the removal of 119 trials across the 41 participants in the onset-match condition, or 2.8% of the data, and 65 trials across the 32 final participants in the full-neighborhood condition, or 2.2% of the data. Since the direction of the effect was predicted in advance, one-tailed tests were used in all the analyses.

Results

As in Experiment 1, all the analyses were performed twice, once including all items, and the second time including only those items with accuracy levels above

80%. The pattern of the analyses was identical in both cases, and the complete analyses are reported below.

We first examined the results for the full-neighborhood condition. There was no difference between the two sets in terms of accuracy [$t_1(31) = -0.58$; $t_2(48) = 0.10$, both $ps > .25$; $\eta^2 = .011$], with average accuracies of 94.9% for items high in neighborhood and 95.4% for items low in neighborhood. There was a difference in the RT data, with slower responses to nonwords high in overall neighborhood than for nonwords low in neighborhood, as is shown in the left panel of Figure 2. This difference was significant by participants, but not by items [$t_1(31) = 2.78$, $p < .005$; $t_2(48) = 1.13$, $p > .10$; $\eta^2 = .200$, a large effect], with average RTs of 509 msec for items high in neighborhood but of 488 msec for items low in neighborhood. The significant effect suggests that nonwords that are overall similar to more real words are responded to more slowly, even when those similar words do not match at word onset.

We then examined the results for the onset-match condition. Again, there was no difference between the two sets in terms of accuracy [$t_1(40) = -0.11$, $p > .25$; $t_2(54) = 0.17$, $p > .25$, with accuracy of scores of 95.1% and 94.8%; $\eta^2 < .001$]. However, unlike with the full-neighborhood set, there was also no effect in the RT data, as is shown in the right panel of Figure 2 [$t_1(40) = 0.46$, $p > .25$; $t_2(54) = 0.47$, $p > .25$; $\eta^2 = .005$], with average RTs of 500 and 501 msec. Apparently, when overall lexical neighborhood was held constant, the number of neighbors that word initially matched the target had no influence on lexical decision responses to nonwords.

These results suggest that neighborhood effects are driven not simply by those words that match the target at onset, but by all the items in the lexicon that are similar to the target. Even when a word is easily distinguishable from the target nonword prior to lexical access, it can still influence the speed at which the target is identified and a lexical decision is made. In fact, this experiment showed no evidence that words that match the target in their initial phoneme have a greater influence on lexical decision than do those that do not match in their initial phoneme. Thus, on the basis of the results of both Experiments 1 and 2, the answer to our question about how to compute a lexical neighborhood is that all words that mismatch in one phoneme should be included in the calculation. The present evidence supports the idea that all of these words influence lexical access. Furthermore, the present data suggest that words that match the target at onset may not play any special role in nonword recognition.

Could these results be an effect of phonotactic probability? The items for both this experiment and Experiment 1 had been selected so as to minimize both phoneme and biphone differences across sets. That said, whereas the phonemes were identical across sets in Experiment 1, they were not quite identical in Experiment 2, and the biphones differed slightly across sets in both experiments. To examine whether these differences resulted in different probabilities across sets, we calculated the frequency with which each phoneme and biphone (sequential pair of pho-

Table 2
Average and Total Frequency-Weighted Neighborhood Density for the Four Sets of Nonwords Used in Experiment 2

Condition	Onset-Matching Neighborhood		Full Neighborhood	
	Average	Total	Average	Total
Onset match				
High set	**11.2**	**314.2**	16.8	472.4
Low set	**6.6**	**183.8**	16.5	460.9
Full neighborhood				
High set	7.2	180.6	**19.9**	**496.3**
Low set	7.4	183.8	**13.3**	**331.3**

Note—The items that differ are in bold.
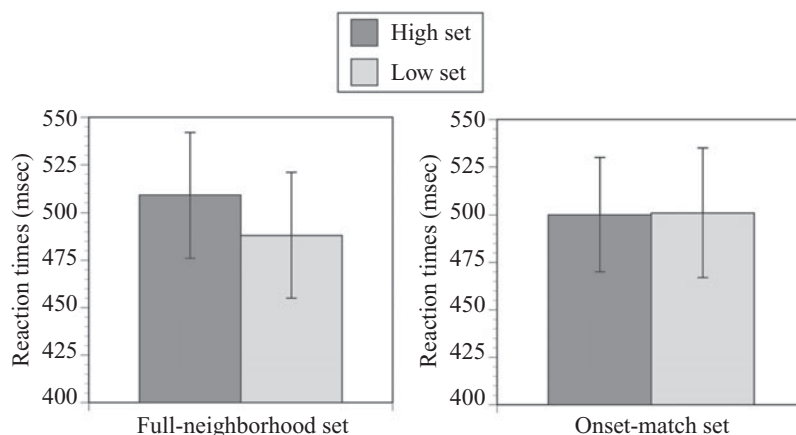
Reaction time data, Experiment 2

Figure 2. Lexical decision reaction times for nonwords in the full-neighborhood condition (left panel) and the onset-match condition (right panel) in Experiment 2. Error bars reflect standard errors.

nemes) occurs in that position in the language and then summed these values to determine a phoneme and biphone probability for each word. These were then summed across all words in each set. The results are shown in Table 3. As is clear from that table, in none of the three sets of words were there any differences in phoneme or biphone frequencies. Clearly, then, the phonotactic probabilities neither of phonemes nor of pairs of phonemes can account for the performance differences found in these two experiments.

Vitevitch (2002) also looked for an effect of onset match. In his study, listeners heard two sets of CVC words, matched for overall neighborhood, but differing in the proportion of neighbors that matched at syllable onset, akin to our onset-match condition. In contrast to our lack of an effect, Vitevitch found a significant difference: Listeners repeated words aloud more quickly and made lexical decisions to them more rapidly when they had fewer neighbors that matched at stimulus onset. His stimuli were real words, whereas ours are nonwords, but there is no theoretical reason why this would result in different effects of lexical neighborhood. Still, that

difference could result in different lexical access patterns. What else can explain this apparent contradiction? One possibility, of course, is that the smaller number of items in this study (28 per set vs. 45 in his study) resulted in the masking of a small benefit for cohorts, one that Vitevitch found. Or perhaps the sets of items in his study had a greater difference in terms of their number of onset-match neighbors, again allowing him to see an effect that was invisible in the present study. A third possibility, however, has to do with the way in which sets of items were balanced. In Vitevitch, items across the two sets had the same initial phonemes, in order to control for differences in voice key sensitivity, but the two sets did not contain the same final consonants and, thus, were not as fully balanced for phonotactics as were the sets of items presented here. Although the two sets did not differ statistically in terms of their phonotactic probabilities, the trend toward a difference was quite close to significance [$t(88) = 1.94$, $p = .055$, two-tailed]. The items in his sparse-onset set, which were responded to more quickly, had slightly higher phonotactic probabilities than did those in the dense-onset set. One possibility, then, is that this difference in phonotactics, rather than a difference in onset-matching neighbors per se, was responsible for the significant effect in his experiment.

EXPERIMENT 3

The previous results suggest that all words similar to a target item become activated, regardless of when in the time course of listening the two items become distinguishable. However, since the difference between sets was not that great, our interpretation is only as accurate as our method of calculating neighborhoods. The one-phoneme substitution metric is only an approximation to the neighborhood weightings performed by the listener. One might expect, for instance, that a neighbor that has more features in common with the target item would

Table 3
Total Phoneme and Biphone Probabilities, Summed Across Items, for the Word Sets Used in Experiments 1 and 2

Condition	Phoneme Frequency	Biphone Frequency
Experiment 1		
Onset match	21.966	20.069
Full neighborhood	21.966	20.079
Experiment 2		
Onset match		
High set	31.065	28.107
Low set	31.115	28.122
Full neighborhood		
High set	27.958	25.117
Low set	27.958	25.113

have a greater influence on perception. The one-phoneme substitution metric does not include these differences; it treats *wed* and *shed* as equally good neighbors of *red*. Although it seems unlikely, it is possible that the neighborhood effect from the previous experiment was due to some difference between the way we calculated neighbors and the way the listener does, rather than to a difference between whether the first phoneme matched or not. We decided to further investigate our interpretation by examining this issue with the phonetic identification task used by Newman et al. (1997). By using phoneme identification with series of syllables, it is possible to select series such that *all* the neighbors differ from the target items in their initial phoneme. In this case, any effect of lexical neighborhood would have to arise from items that mismatched in their initial phonemes. This would provide additional evidence that even when a word is distinguishable from a target at syllable onset, the word still becomes partly activated during the course of processing.

The CVC syllables used by Newman et al. (1997) and Ganong (1980) consisted of a VOT continuum for the initial phoneme, a shared vowel, and a final consonant. Thus, the same vowel was paired with all four items. The result of this was that any real word that differed from the target in either the initial consonant or the final consonant would necessarily be a neighbor to two of the four endpoints. This item, then, would not contribute to the differential neighborhood effect. As an example, one set used by Newman et al. (1997) involved a VOT continuum ranging from /b/ to /p/, followed by the vowel /e/ and the consonants /ʃ/ or /θ/. This resulted in four nonsense syllables: *beysh*, *peysh*, *beyth*, and *peyth*. For this pair of series, the items *beyth* and *peysh* are higher and *beysh* and *peyth* are lower in frequency-weighted neighborhood density. The real word *bake* is an example of an item that differs in the final consonant. This item is necessarily a neighbor to both *beyth* and *beysh*. Thus, it cannot contribute to the differential bias toward the /b/ in one series and the /p/ in the other. Similarly, the word *faith*, which differs in its initial consonant, is a neighbor to both *beyth* and *peyth* and cannot contribute to a bias for either initial phoneme.

The real neighborhood bias for these items, then, was driven by those syllables that matched both the initial and the final consonants and mismatched in the vowel (the one phoneme common to all four syllables). For example, both of the words *bath* and *push* are neighbors to only one of the four endpoints (*beyth* and *peysh*, respectively). It is these types of neighbors that give those nonwords their relatively higher neighborhood frequency. As a consequence, the computation of neighborhoods for these syllables will yield similar results for a neighborhood rule and a cohort rule. As was noted by Newman et al. (1997), their results could not distinguish between these two alternatives.

Distinguishing between a neighborhood-based rule for computing similarity to real words and a cohort-based rule requires that the phoneme that is constant across the two series be the initial phoneme. In such a pair of series, any word that mismatches one of the series' endpoints in the initial position will be a neighbor to only that endpoint. Thus, only words that are not a part of the cohort to the target will contribute to differential neighborhood size. In order to select pairs of series for which only neighbors that differ in the initial consonant can contribute to a neighborhood effect, it is not possible to select target items that share the same vowel. Rather, the vowel has to be one of the phonemes that differ across series. This poses a potential problem of coarticulation. In previous studies, phoneme identification has been examined when the test continuum is syllable initial (Ganong, 1980; Pitt & Samuel, 1993) and when it is syllable final (McQueen, 1991). In all cases, the vowel remained constant. This meant that the final consonant from one item could be spliced onto another item with less worry that there might be coarticulated acoustic qualities from the neighboring vowel that could cause misinterpretation when the context changed.

To help address this issue, we decided to create two sets of series. One pair of series consisted of an initial-consonant continuum, followed by differing vowels, and a shared final consonant. Neighborhood effects for these items will be driven primarily by items differing in their final consonant from the target item. By all accounts, these items should produce neighborhood effects. After all, it is only at the end of the syllable for these items, when activation is well under way, that the nonword becomes distinct from its word neighbors. This is the prototypical situation in which one would expect a neighborhood effect, and nearly all theories of word recognition would predict activation of real-word neighbors in this situation. If the use of series with two different vowels causes problems in this task, we should be able to detect that here, where neighborhood effects are clearly expected.

The second pair of series consisted of a shared initial consonant, followed by differing vowels, and a syllable-final consonant continuum. This pair provides a test case for the role of neighborhood items that differ from the target in their initial consonant. Since all four syllables share the initial consonant, items differing syllable initially will drive any neighborhood effect for these series. In addition, we specifically selected series in such a way that none of the four endpoints had *any* neighbors with the same initial consonant as the target items. Thus, there is no way a neighborhood (cohort) effect could arise from items that matched the target syllable initially.

Both series were based on a /ʃ/–/tʃ/ continuum. The first pair of series ranged from /ʃoɪf/ to /tʃoɪf/ (*shoif* to *choif*) and from /ʃof/ to /tʃof/ (*shof* to *chof*). These items all share the final consonant, so words differing in that consonant (e.g., *choice*) will be neighbors to only one endpoint item and should drive the neighborhood effect. The second pair of series was designed to test the hypothesis that neighbors that do *not* overlap on the initial consonant contribute to neighborhood density effects on phoneme perception. They ranged from /zʊʃ/ to /zʊtʃ/ (*zush* to *zuch*, with the vowel sound in *bush*, not *zoo*) and from /zɛʃ/ to /zɛtʃ/ (*zehsh* to *zehch*). These items share

an initial consonant, so items differing in that consonant (such as *bush*) will be neighbors to only one item. In fact, these items were chosen such that all of their real-word neighbors differed from the target syllables in their initial consonant.

No previous study of neighborhood effects on phoneme identification has used a contrast in syllable-final position. However, two studies of the lexical effect have used a syllable-final contrast (McQueen, 1991; Pitt & Samuel, 1993). In general, the effect of one end of a continuum being a word while the other is not (e.g., the *fish–fiss* and *kish–kiss* series used by McQueen, 1991) is small and tends to appear most robustly in the faster responses made by listeners. In Experiment 3, all of the items were nonwords. Consequently, only lexical neighborhood could influence the perception of the final phoneme. Thus, this experiment, if anything, was likely to show effects even smaller than the previous lexical effects (where both lexical status and neighborhood tended to favor one end of each series). In addition, on the basis of the results of Newman et al. (1997), the effects of lexical neighborhood are transient. Newman, Sawusch, Luce, and Aubin (2000) have also shown that neighborhood effects emerge earlier in the time course of processing than do effects of lexical status. Consequently, the predictions based on the neighborhood approach are that the effect will be small and will tend to show up in the fastest responses made by the listeners.

Method

Participants. The participants were 29 members of the University of Iowa community, who took part in this experiment for course credit. Every participant listened to the syllables from both pairs of series. All were native speakers of English, with no reported history of a speech or hearing disorder. The data from 1 participant were removed from analysis (both sets of series) when it was determined that she was not a native speaker of English. Data from an additional 6 participants for the *shoif–chof* series and 5 participants for the *zush–zech* series were dropped from analysis for failure to classify the two endpoints of one series at 80% or better. In addition, data from an additional 2 participants for the *shoif–chof* series were dropped for equipment failure/experimenter error, and data from an additional participant for the *zush–zech* series were dropped because of a low overall response rate. This left a total of 20 participants in the *shoif–chof* condition and 22 participants in the *zush–zech* condition.

Stimuli. Although all consonants are coarticulated with the adjacent vowel, some consonants are easier to segment acoustically, in the temporal and spectral domain, from the adjacent vowel. We selected series so that all four consonants were fricatives or affricates. The presence of noise in the waveform for these consonants makes the process of identifying reliable acoustic criteria for editing easier. We also selected items so that the same two phonemes were used for the continua in both series.

The sets of items ranged from /ʃɔɪf/ to /tʃɔɪf/ (*shoif* to *choif*) and from /ʃof/ to /tʃof/ (*shohf* to *chohf*) for those items with the continuum at the onset. The series with the consonant continuum at offset ranged from /zɛʃ/ to /zɛtʃ/ (*zehsh* to *zehch*) and from /zʊʃ/ to /zʊtʃ/ (*zush* to *zuch*, with the vowel in *bush*, not *boot*). For convenience, we will refer to each pair of series by using an orthographic spelling of the endpoints with the higher frequency-weighted neighborhood density. According to this notation, our series were *shof–choif* and *zush–zech*. Table 4 lists the neighbors and the total neighbor-

hood values for each of these eight items. Clearly, any effect of neighborhoods in the *zush–zech* set can be caused only by words that do not match at stimulus onset.

A female native speaker of English (R.S.N.) recorded the syllables /ʃɔɪf/, /ʃof/, /zʊʃ/, /zɛʃ/, and /zʊtʃ/, using a Shure SM81 cardioid microphone. All of the tokens were recorded at a 44.1-kHz sampling rate with 16-bit quantization and were stored on computer disk.

The vowel in /ʃof/ was significantly shorter than that in /ʃɔɪf/, so every other pitch pulse was reduplicated until their durations were nearly identical (239.4 msec for /o/, 240.2 msec for /ɔɪ/). The /f/ frication from /ʃof/ was removed and was replaced with the corresponding frication from /ʃɔɪf/. This produced two VC syllables with similar duration and an identical final consonant.

The /ʃ/ from /ʃɔɪf/ was removed and used as the base *sh* frication. A series ranging from /ʃ/ to /tʃ/ was created by removing successive sections from the onset of the frication. Cuts were made at zero crossings; because of this, the durations of removed sections were not entirely equal but averaged 10 msec each. This resulted in an

Table 4
Computation of Frequency-Weighted Neighborhood Density for Endpoint Items in Experiment 3

Word	Frequency	Log_{10} (Frequency $\times$ 10)	Familiarity
Target: Shoif (/ʃɔɪf/) Neighborhood			
Chef	9	1.95	6.75
$\Sigma \log_{10}$ (frequency $\times$ 10) =		1.95	
Target: Choif (/tʃɔɪf/) Neighborhood			
Choice	113	3.05	6.92
Chief	119	3.08	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		6.13	
Target: Shohf (/ʃof/) Neighborhood			
Chef	9	1.95	6.75
Loaf	4	1.60	7.00
Show	287	3.46	7.00
Shown	171	3.23	6.83
Shore	61	2.79	7.00
Showy	1	1.00	6.42
Chauffeur	4	1.60	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		13.63	
Target: Chohf (/tʃof/) Neighborhood			
Choke	9	1.95	7.00
Chief	119	3.08	7.00
Chore	7	1.85	6.75
Chose	37	2.57	6.50
Loaf	4	1.60	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		11.05	
Target: Zush (/zʊʃ/) Neighborhood			
Bush	15	2.18	7.00
Push	37	2.57	6.92
$\Sigma \log_{10}$ (frequency $\times$ 10) =		4.75	
Target: Zuch (/zʊtʃ/) Neighborhood			
No neighbors			
$\Sigma \log_{10}$ (frequency $\times$ 10) =		0.00	
Target: Zehsh (/zɛʃ/) Neighborhood			
Mesh	4	1.60	6.58
$\Sigma \log_{10}$ (frequency $\times$ 10) =		1.60	
Target: Zehch (/zɛtʃ/) Neighborhood			
Etch	1	1.00	6.33
Fetch	6	1.78	7.00
Wretch	2	1.30	6.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		4.08	

11-member series with frication duration ranging from 163.6 msec (*sh*) to 63.0 msec (*ch*). Each of these 11 frication durations was spliced to the beginning of both /oʃ/ and /oɪʃ/ to create the two series.

For the *zush–zech* series, the /z/ from /zɛʃ/ was removed and was replaced with the corresponding frication from /zʊʃ/. One pitch pulse was removed from the /ɛ/ in order to make the item sound more natural. Although the two vowels were not identical in duration, they were quite close (/ʊ/, 145.2 msec; /ɛ/, 158.0 msec), and no change was made to correct for this slight difference. The duration of the natural /ʃ/ frication and /tʃ/ closure + frication were also quite similar (/ʃ/, 202.4 msec; /tʃ/, 186.9 msec).

To make the syllable-final /ʃ/–/tʃ/ continuum, the portions representing the two final consonants were blended together in varying amplitude ratios. For the most extreme *ch* endpoint, the frication from the /ʃ/ was reduced in amplitude by 18 dB before being blended with the /tʃ/ closure + frication. For the next item in the series, the /ʃ/ was reduced by 15 dB before blending. This continued in 3-dB steps until the 2 items were blended at equal intensity. An additional 3 items beyond that point were created by reducing the amplitude of the /tʃ/ by successive 3-dB steps prior to blending with the full-amplitude /ʃ/. This resulted in a 10-item series ranging from a good *sh* to a good *ch*. Each of these 10 blends was spliced to the end of the /zʊ/ and the /zɛ/ to create the two 10-item series.

Procedure. The listeners were tested individually. Each heard both pairs of series in a blocked format. Half the participants heard the *shof–choif* series first, followed by the *zush–zech* series, whereas the remaining participants heard the series in the reverse order. Stimulus presentation and response collection were controlled by a Macintosh 7100/AV computer. The stimuli, which were stored on disk, were presented binaurally through AudioTechnica ATH-M40fs headphones. The syllables making up each pair of series were presented in random order. The listeners were asked to identify the initial (or final) phoneme as *sh* or *ch* as quickly and accurately as possible by pressing one of two buttons on a computer-controlled response box. The mapping of response to hand was counterbalanced across listeners.

The presentation pace depended on the listeners' response speed. The next trial began 1 sec after the listener had responded or after an interval of 3 sec from stimulus onset had elapsed, whichever came first. The listeners' responses and RTs were recorded for each stimulus. Responses from the first block of trials with each series (two repetitions of each item) were considered practice and were not included in subsequent data analysis. After the practice block, the stimuli were presented in blocks of either 60 or 66 trials (three repetitions of each of the 20 or 22 items). All the listeners participated in five blocks of experimental trials for each pair of series, which resulted in a total of 15 responses to each stimulus. RTs greater than 1,500 msec from the start of the stimulus were removed from analysis for the *shof–choif* series; for the *zush–zech* series, the critical phoneme was at the end of the stimulus, requiring that the listeners wait longer before responding. Since the initial CV of these items was 274 msec for /zʊ/ and 287 msec for /zɛ/, listeners needed to wait approximately an additional 300 msec before responding to items in this series than in the *shof–choif* series. In order to make our cutoffs comparable across series, we chose 1,800 msec as the cutoff for the *zush–zech* series. Any response greater than 1,800 msec from the start of the stimulus was removed from analysis. This resulted in the loss of 126 out of 6,563 trials with responses in the *shof–choif* series (or 1.9%) and the loss of 46 out of 6,520 trials in the *zush–zech* series (or 0.7%).

Results and Discussion

First, the overall data for each pair of series were examined. The percentage of *sh* responses was determined for each stimulus in each series for every listener. Any response with an RT greater than 1,500 msec was eliminated. The data from each listener were examined for consistency in classifying the endpoints of the series. Any listener who could not classify the endpoints consistently (at 80% accuracy or better) was dropped from the analysis of that pair of series, as has been described previously.

The previous literature on these neighborhood effects has included two types of measures of perceptual change, rather than just one: shifts in category boundary and changes in percentage of responses to one phoneme (here, the *sh*). Whereas the category boundary movement should indicate changes in perception of ambiguous stimuli, the overall percentage data should be sensitive to changes away from the boundary, as well as those at the boundary, and thus may be more sensitive (see Pitt & Samuel, 1993). We will report percentage data, but in order to maintain comparability with the prior literature, category boundary data were also calculated and are included as notes. The category boundary for each listener was determined for each series by linear interpolation between the two stimuli on either side of the boundary; the total percentage of *sh* responses was calculated by totaling the *sh* responses by each listener to all of the stimuli in each series. We conducted paired *t* tests on the data on the basis of each measure. Because we were predicting an effect in a specific direction (more responses for the nonword with a higher neighborhood frequency), we used one-tailed tests. For all analyses, η^2 values are provided.

We then partitioned the data for each listener into three subsets based on RT. This partitioning has been done frequently in the prior literature and has been shown to be especially sensitive to small transient effects, such as those of neighborhoods (Fox, 1984; Miller & Dexter, 1988; Newman et al., 1997; Pitt & Samuel, 1993). Partitioning in this manner seems to be the most appropriate way to examine transient temporal effects of neighborhood activation on phoneme perception. For each listener's responses to each stimulus, we determined the mean and standard deviation and then translated each individual RT into a *z* score. Scores of 0.43 and −0.43 divide the distribution into three approximately equal portions. Any RT equivalent to a *z* score less than −0.43 was considered fast, any RT greater than a *z* score of 0.43 was considered slow, and any RT *z* score between these values was considered intermediate. Once each listener's data had been partitioned, identification functions, category boundaries, and percentages of *sh* responses were determined for each series in each partition, as was described previously for the overall analysis. We used 2 two-way ANOVAs with RT partition as one factor and series as the other to examine the percentage of *sh* responses and the category boundary data. These were followed by paired *t* tests (planned comparisons) in each RT partition, as was described previously for the overall analysis. This set of analyses is similar to that used by Newman et al. (1997) for their studies of neighborhood influences on phoneme perception.

We first examined the data from the *shof–choif* series. Overall, there was a significant change in the percentage of *sh* responses [$t(19) = 3.82$, $p < .005$; $\eta^2 = .43$, indicating a large effect]. The listeners gave more *sh* re-

sponses to the *shof–chof* series and more *ch* responses to the *shoif–choif* series. This difference is consistent with the difference in neighborhood frequency. That is, the listeners gave more responses to each series that were consistent with the endpoint that had the higher frequency-weighted neighborhood density (*shof* and *choif*). Figure 3 displays the mean identification functions for these two series across listeners and shows a change in the locus of the category boundary for the two series.

The results of partitioning the RT data are shown in Figure 4, with identification functions for the fast responses on the top and slow responses on the bottom. A two-way ANOVA for the percentage of *sh* responses showed a marginal effect of RT partition [$F(2,38) = 3.09, p < .10; \eta^2 = 0.024$, or a small effect], a significant effects of series [$F(1,19) = 10.63, p < .005; \eta^2 = .089$, indicating an intermediate effect], and no interaction ($F < 1; \eta^2 = .004$). These results suggest that there was a consistent effect of neighborhood across all three speed ranges and that there was also a trend toward fewer *sh* responses at longer RTs.

The consistent neighborhood effect was also demonstrated in the planned comparisons. There was a significant effect in all three RT ranges [for the fast responses, $t(19) = 2.06, p < .05$; for the intermediate responses, $t(19) = 3.44, p < .005$; and for the slow responses, $t(19) = 1.91, p < .05$]. The effect sizes were large in all three speed ranges, with η^2 values of .183, .384, and .161, respectively.[2]

These results show that neighborhood effects can be found for pairs of series that do not have the same vowel. Thus, potential differences in coarticulation among these items were not sufficient to prevent or mask the effects of lexical activation of similar words. Therefore, these
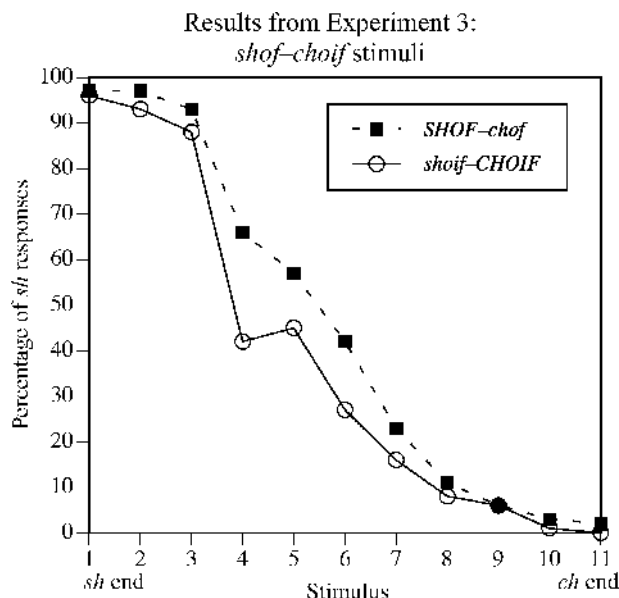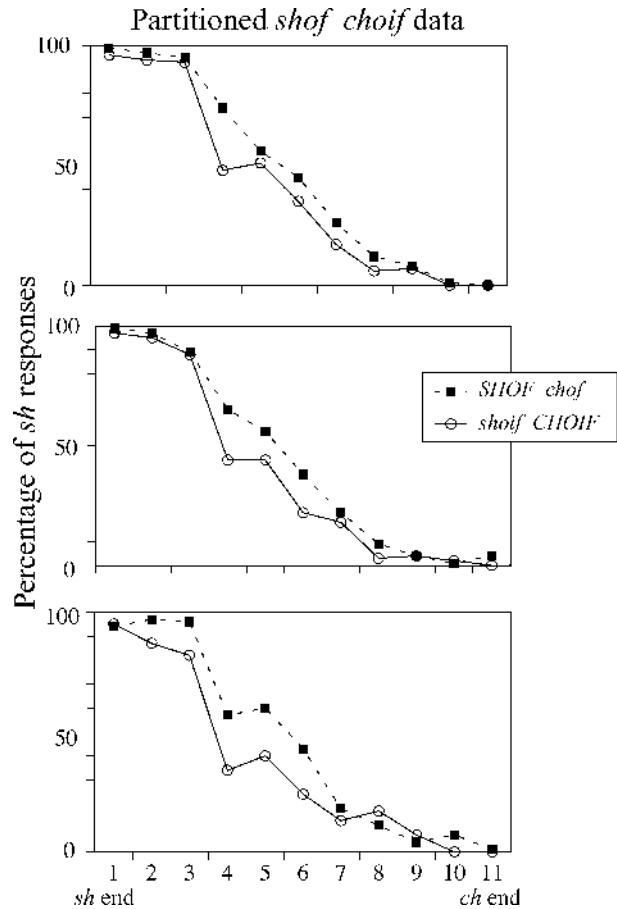

Figure 4. Group identification functions for the *shoif–choif* and *shof–chof* series in three reaction time partitions: fast, intermediate, and slow.

Figure 3. Group identification functions for the *shoif–choif* and *shof–chof* series.

types of series can be used to examine the role of neighbors that do not match word initially.

The data for the *zush–zech* series were examined in the same manner. Figure 5 shows the overall labeling results, and Figure 6 shows the results across the three RT partitions. In the overall data, there was a significant change in the percentage of *sh* responses [$t(21) = 2.00, p < .05$], with η^2 values indicating a large effect size ($\eta^2 = .16$). We then partitioned on the basis of RT and found significant but small effects for partition [$F(2,42) = 4.31, p < .01; \eta^2 = .041$] and neighborhood [$F(1,21) = 4.43, p < .05; \eta^2 = .013$] and no interaction ($F < 1; \eta^2 = .005$). The planned comparisons showed only marginal effects of neighborhood for both the fast responses [$t(21) = 1.59, p < .07$] and the intermediate responses [$t(21) = 1.48, p < .08$] and no effect for the slow responses [$t(21) = 0.16, p > .10$].[3] Despite the fact that the effects for the fast and intermediate partitions did not reach statistical significance, they were of intermediate effect size ($\eta^2 = .11, .09$, and .001 for the three partitions, respectively).

These results demonstrate that there may be a small but somewhat variable effect of neighborhood for the *zush–zech* series, but it is clearly less robust than neigh-
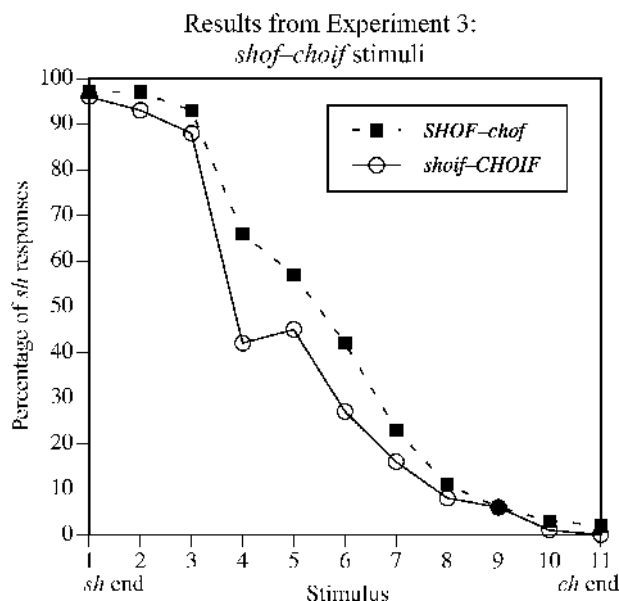
Results from Experiment 3:
shof–choif stimuli

Legend:
- ■ · SHOF–*chof*
- —○— *shoif*–CHOIF

Figure 5. Group identification functions for the *zush–zuch* and *zesh–zech* series.

borhood effects for the *shof–choif* series. Since none of the endpoints in the *zush–zech* series have any neighbors that match on their initial phoneme, any effect that does occur must be caused by neighbors such as *push*, *mesh*, and *fetch*, which do not match syllable initially. This suggests that these neighbors may be activated as part of lexical access, even when the syllables were presented in isolation (where the onset of the nonword was not in question). If so, it would support the findings from Experiment 2, which also showed that all neighbors, regardless of the location of their difference, receive activation when a target is perceived. Yet the effect, although significant, is clearly quite weak.

There are many possible reasons for the difference in the size of the effect across the two conditions. One obvious explanation is that this might be an indication of the time course of word recognition. Although items that mismatch word initially may be part of the lexical neighborhood, they may play a lesser role than do words that mismatch at the end of the stimulus, at least at the point in time at which our listeners' responded. Were this the case, we might expect that the breakdown of responses into RT partitions would demonstrate a different pattern of responses for the two series. Unfortunately, since the vowels are not the same across all four series and the time course of processing a /tʃ/–/ʃ/ contrast in initial and final positions may not be the same, comparisons of time course across the two sets are tenuous. McQueen (1991) has shown that word-final contrasts tend to demonstrate lexical effects more strongly with short RTs, whereas effects with word-initial contrasts tend to appear in later RT partitions. Our results for the *zush–zech* neighborhood series appear to be consistent with this pattern of findings, with a significant effect of neighborhood for

the fast and intermediate speed responses. This suggests that the temporal course of lexical access may indeed cause different patterns of effects for word-initial contrasts than for word-final ones. Furthermore, to the extent that neighborhood effects dissipate more rapidly than lexical status (word or not) effects, the small effect that was found in the percentage of response data for *zush–zech* may represent an effect that is already fading or dissipating. This is certainly consistent with the small effect that we observed.

A second possible explanation has to do with the difference between the continua in the two series. Since the /ʃ/–/tʃ/ continua are not identical, the step size may be perceptually different across the two series. We do not expect that the spacing between members of the two series are necessarily identical, either acoustically or perceptually. Perhaps there were more ambiguous items in the initial /ʃ/–/tʃ/ continuum than in the final /ʃ/–/tʃ/ continuum. To assess the degree to which the series did or did not contain ambiguous stimuli near the boundary, we fit a psychometric function (an ogive) to the data for each of the two series for each listener, according to the
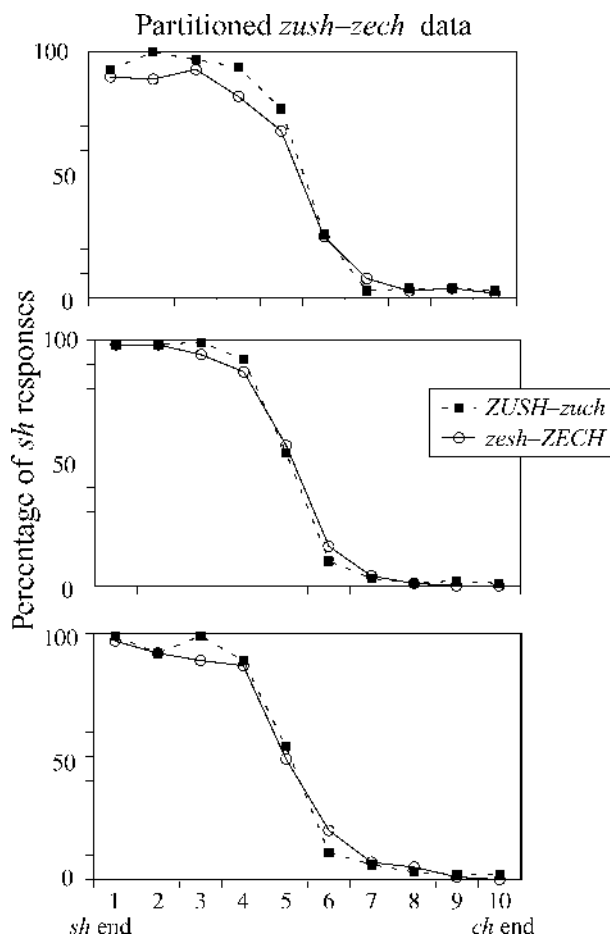
Partitioned *zush–zech* data

Legend:
- ■ · ZUSH–*zuch*
- —○— *zesh*–ZECH

Figure 6. Group identification functions for the *zush–zuch* and *zesh–zech* series in three reaction time partitions: fast, intermediate, and slow.

method described by Engen (1971). For each psychometric function, a slope was derived. High slopes correspond to more steplike classification functions (ones with less ambiguity), whereas low slopes correspond to more gradual classification functions, indicating greater ambiguity in the listener's perception of stimuli in this series. The slopes for the two series did differ significantly [$F(1,40) = 63.58$, $p < .0001$], with average slopes of 1.83 for the *zush–zech* series and 0.77 for the *shof–choif* series. The slope difference implies that the locus of the category boundary is more ambiguous in the *shof–choif* series. This difference is likely to result in a larger boundary shift for the *shof–choif* series, which is exactly what we found. Thus, the weaker results for the *zush–zech* series may simply be caused by differences in step size across the series. There is no unambiguous evidence for a difference in the size of the neighborhood influence across the two series.

Despite this fact, the effect size for the *zush–zech* series clearly appears to be somewhat small, and a stronger, more consistent effect would clearly be more convincing. To that end, Experiment 4 was an attempt to replicate these results with another series.

EXPERIMENT 4

The *zush–zech* series from Experiment 3 showed only very slight effects of lexical neighborhood. This may, in part, have been a result of the fact that the neighborhood sizes for these words were quite small. Regardless of the reason, the findings would be more convincing if they could be replicated. To help address this issue, another pair of series were created in which neighborhood effects would be driven by items differing from the target item in their onset (initial consonant). As in Experiment 3, none of the four endpoints had *any* neighbors that shared the initial consonant. In terms of our computations, all four endpoints had equivalent, empty (zero-sized) cohorts. Thus, it was not possible for neighborhood effects to arise from items that matched the target syllable initially. More critically, the neighborhood size was much greater for these series than for those in Experiment 3. Whereas the endpoints in Experiment 3 had from zero to three neighbors (and neighborhood values of 0.0–4.75), those in the present series ranged from three to seven neighbors (with neighborhood values of 7.49–20.56). Thus, if the weakness of the effect in Experiment 3 was a result of the relatively small neighborhood sizes and differences in neighborhood sizes for these nonwords, we would expect to find a much stronger effect in the present experiment.

The series used ranged from /zʊk/ to /zʊt/ (*zook* to *zoot*, with the vowel sound in *book*, not *zoo*) and from /zɔk/ to /zɔt/ (*zawk* to *zawt*). These items share an initial consonant, so items differing in that consonant (such as *book*) will be neighbors to only one item. As in Experiment 3, we divided the participants' responses into three RT partitions and expected the results to appear most strongly in the fast (or possibly, fast and interme-

diate) responses. We did not expect to find any effect of neighborhood in the listeners' slow responses.

Method

Participants. The participants were 34 members of the University of Maryland community, who took part in this experiment for a cash payment. All were native speakers of English, with no reported history of a speech or hearing disorder. The data from 3 participants were excluded from analysis when it was determined that they were not native speakers of English. The data from another 3 participants were excluded after they reported having been diagnosed with an attention disorder. Data from 3 participants were lost as a result of equipment failure. Finally, data from 1 participant were dropped from the analysis for failure to classify the two endpoints of one series at 80% or better. This left a total of 24 participants in the final analysis.

Stimuli. The sets of items ranged from /zʊk/ to /zʊt/ (*zook* to *zoot*) and from /zɔk/ to /zɔt/ (*zawk* to *zawt*). Using the same method of notation as that in Experiment 3, we will refer to this series as *zook–zawt*. Table 5 lists the neighbors and the total neighborhood values for each of these four items. Clearly, any effect of neighborhoods can be caused only by words that do not match at stimulus onset.

A female native speaker of English (R.S.N.) recorded the four endpoint syllables, using a Shure SM81 cardioid microphone. All of the tokens were recorded at a 44.1-kHz sampling rate with 16-bit quantization and were stored on computer disk.

The /z/ from /zɔt/ was removed and was replaced with the corresponding frication from /zʊt/. The first few pitch pulses of the /ɔ/ vowel were amplified slightly so as to better blend with the /z/ from /zʊt/. The /ɔ/ vowel is inherently longer than the /ʊ/; in these recordings, the /ɔ/ was 198 msec, and the /ʊ/ was 103 msec. Since there was no reason to expect durational contrast effects to influence a /t/–/k/ series, this vowel duration difference was maintained in the final stimuli. The transitions from the vowel to the final consonant for /zʊt/ and /zʊk/ and for /zɔt/ and /zɔk/ were determined by inspection of spectrograms for the tokens. These transitions were removed and blended in equal ratios. That is, the transitions from /ʊ/ to /t/ and to /k/ were blended to create a composite, so that the transitions from the vowel contained acoustic cues to both final consonants. Similarly, the transitions from /ɔ/ to /t/ and to /k/ were blended. These transition periods were 54 msec for the /zɔk/–/zɔt/ series and 36 msec for the /zʊk/–/zʊt/ series. After blending, the composite was reattached to the /zɔ/ and /zʊ/.

To make the syllable-final /t/–/k/ continuum, the portions representing the two final consonants from /zʊk/–/zʊt/ were blended together in varying amplitude ratios, with step sizes of 3 dB. First, the /k/ burst was amplified to be similar in RMS amplitude to that of the /t/. The two were then mixed together at equal amplitude levels. To a listener, this item sounded relatively /k/-like, but not completely so. Two more items were made for the /k/ end of the continuum, with the /k/ amplitude being 3 and 6 dB more intense than the /t/ amplitude before blending. An additional 8 items were made toward the /t/ end of the continuum, with the /k/ amplitude decreased in 3-dB steps relative to the /t/ amplitude. This resulted in an 11-item series ranging from a good *k* to a good *t*.

Procedure. The listeners were tested individually and heard all 22 items in random order. Stimulus presentation and response collection were controlled by a Macintosh 7100/AV computer. The stimuli, which were stored on disk, were presented binaurally through AudioTechnica ATH-M40fs headphones. The listeners were asked to identify the final phoneme as *t* or *k* as quickly and accurately as possible by pressing one of two buttons on a computer-controlled response box. The mapping of response to hand was counterbalanced across listeners.

The presentation pace depended on the listeners' response speed. The next trial began 1 sec after the listener had responded or after an interval of 3 sec from stimulus onset had elapsed, whichever

Table 5
Computation of Frequency-Weighted Neighborhood Density
for Endpoint Items in Experiment 4

Word	Frequency	$\log_{10}$ (Frequency $\times$ 10)	Familiarity
Target: Zook (/zʊk/) Neighborhood			
Book	193	3.29	6.91
Cook	47	2.67	7.00
Hook	5	1.70	6.75
Look	399	3.60	7.00
Shook	57	2.76	6.75
Took	426	3.63	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		17.64	
Target: Zoot (/zʊt/) Neighborhood			
Foot	70	2.85	7.00
Put	437	3.64	7.00
Soot	1	1.00	6.58
$\Sigma \log_{10}$ (frequency $\times$ 10) =		7.49	
Target: Zawk (/zɔk/) Neighborhood			
Balk	1	1.00	6.18
Chalk	3	1.47	7.00
Gawk	1	1.00	6.25
Hock	14	2.15	7.00
Talk	154	3.19	7.00
Walk	100	3.00	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		11.81	
Target: Zawt (/zɔt/) Neighborhood			
Aught	68	2.83	7.00
Bought	56	2.75	7.00
Caught	98	2.99	7.00
Fought	46	2.66	7.00
Sought	55	2.74	6.58
Taught	66	2.82	7.00
Thought	515	3.71	7.00
$\Sigma \log_{10}$ (frequency $\times$ 10) =		20.56	

came first. The listeners' response and RT were recorded for each stimulus. Responses from the first block of trials with each series (two repetitions of each item) were considered practice and were not included in subsequent data analysis. After the practice block, the stimuli were presented in blocks of 66 trials (three repetitions of each of the 22 items). All the listeners participated in five blocks of experimental trials for the pair of series, which resulted in a total of 15 responses to each stimulus.[4] As with the final target series from Experiment 3, responses more than 1,800 msec from the start of the trial were excluded; this resulted in a loss of 74 trials out of 7,737, or 1.0%.

Results and Discussion

First, the overall data for the two series were examined. The percentage of *k* responses was determined for each stimulus in each series for every listener. Any response with an RT greater than 1,800 msec was eliminated. The data from each listener were examined for consistency in classifying the endpoints of the series. Any listener who could not classify the endpoints consistently (at 80% accuracy or better) was dropped from the analysis of that pair of series, as was described previously. As in Experiment 3, the category boundary for each listener was determined for each series by linear interpolation between the two stimuli on either side of the boundary, and the total percentage of *k* responses given by each listener to all of the stimuli in each series was

tabulated. We conducted paired *t* tests on the overall data and then partitioned the data for each listener into three subsets based on RT.

Overall, there was a significant change in the percentage of *k* responses [$t(23) = 5.25$, $p < .0001$; $\eta^2 = .545$, a large effect]. The listeners gave more *k* responses to the /zʊk/–/zʊt/ series, and more *t* responses to the /zɔk/–/zɔt/ series, suggesting that they labeled ambiguous items in a manner consistent with the difference in neighborhood frequency. The overall labeling data for the two series are shown in Figure 7, where a clear difference in labeling can be seen for the two series.

The results of partitioning the RT data are shown in Figure 8, with the fast partition on the top. The two-way ANOVA on the percentage of *k* responses data showed significant overall effects for both partition [$F(2,46) = 4.10$, $p < .05$] and neighborhood [$F(1,23) = 23.48$, $p < .0005$], as well as a significant interaction [$F(2,46) = 15.52$, $p < .0001$]. Size-of-effect indicators suggest that while the effect of partition was small ($\eta^2 = .029$), the effects of both neighborhood and the interaction were intermediate ($\eta^2 = .082$ and .088, respectively). Clearly, then, the effect of neighborhood was much stronger here than in Experiment 3. The planned comparisons showed significant effects of neighborhood for the fast and intermediate speed responses, but no effects for the slow partition [for fast responses, $t(23) = 5.96$, $p < .0001$; $\eta^2 = .607$; for intermediate responses, $t(23) = 3.09$, $p < .005$; $\eta^2 = .293$; and for slow responses $t(23) = -0.78$, $p > .20$; $\eta^2 = .026$].[5]

These results demonstrate a significant effect of neighborhood for the *zook–zawt* series. Since none of the endpoints in the *zook–zawt* series have any neighbors that match on their initial phoneme, the effect must be caused by neighbors, such as *book* and *thought*, that do not match

Results from Experiment 4:
zook–zawt stimuli
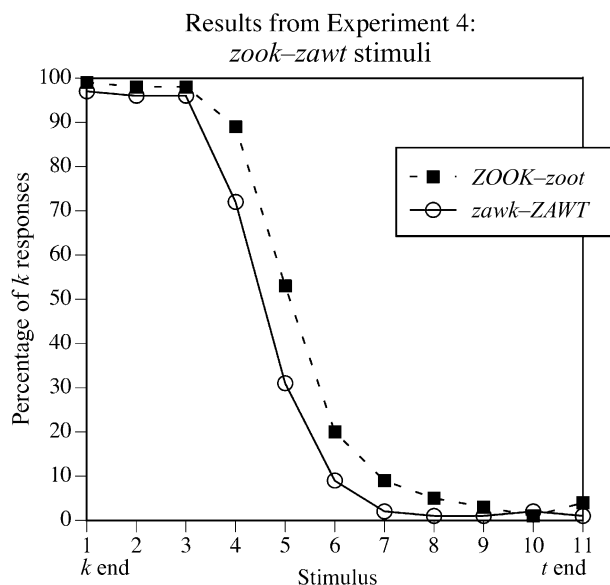
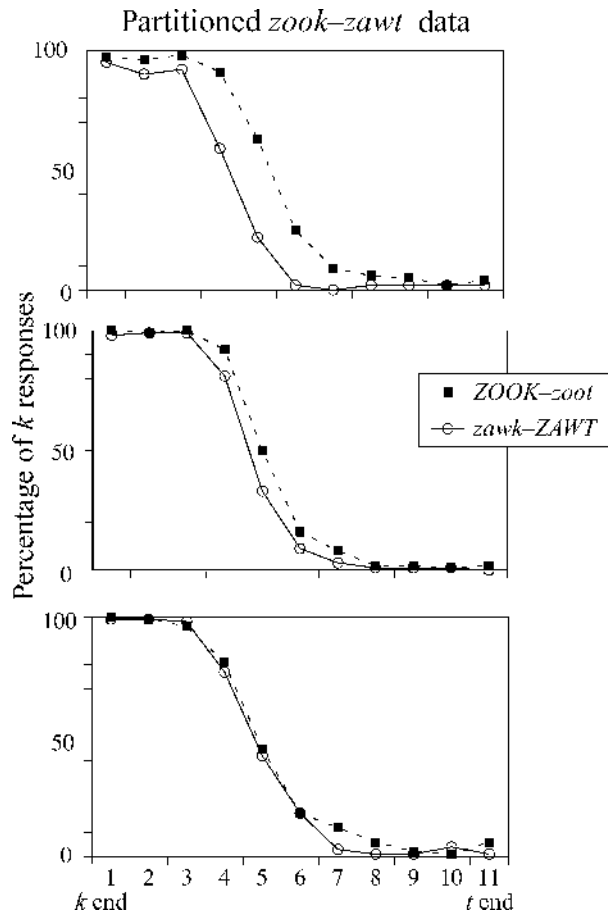**Figure 7. Group identification functions for the *zook–zoot* and *zawk–zawt* series.**

Figure 8. Group identification functions for the *zook–zoot* and *zawk–zawt* series in three reaction time partitions: fast, intermediate, and slow.

syllable initially. This suggests that these neighbors are activated as part of lexical access, even when the nonwords (or words) are presented in isolation (where the onset of the target was not in question). Moreover, unlike the effects in Experiment 3, which were somewhat tenuous, the results here are quite clear. Apparently, all neighbors, regardless of the location of their difference, receive activation when a target is perceived.

There are several possible explanations for the difference in the size of the effect across Experiments 3 and 4. One possibility is that the /t/–/k/ series had more ambiguous items than did the /ʃ/–/tʃ/ series from the prior experiment. To determine this, we examined the slopes of the psychometric functions for this series, as we did for the series in Experiment 3. The slope for the *zook–zawt* series were very similar to that for the previous *zush–zech* series [1.87, as compared with 1.83; $F(1,44) < 1$]. This suggests that the difference is not a result of the present series containing more ambiguous items. Instead, the most likely reason is that the items in the present experiment had larger neighborhood differences between endpoints (as well as larger overall neighborhoods). Apparently, the weakness of the effect in Experiment 3 was not

a result of the fact that the neighbors were all mismatching initially; it was, instead, a result of the relative paucity of those neighbors.

Could the effects in these three series (*shoif–chof*, *zush–zech*, and *zook–zawt*) actually have been the result of coarticulation from the vowel, and not from lexical neighborhoods? Perhaps, but that seems unlikely. Although the change in vowel across series does lead to the possibility of coarticulatory influences (despite the use of fricatives and affricates), there is no reason to believe that any such coarticulatory effects would be in the same direction as neighborhood effects in all three pairs of series. Had the shifts in category boundaries across series been a result of the change in vowel identity, we would expect that the pattern would, by chance, have gone in the opposite direction of neighborhoods for at least one of the three series. Still, it is not entirely possible to eliminate the possibility of coarticulatory effects when vowel identity is altered.

Unfortunately, it is not possible to test for such coarticulatory effects directly. Although we could present the series without the critical consonant, thus eliminating the predicted neighborhood effects (while presumably maintaining any coarticulatory effect), this would lead to the series /ʃɔɪ/–/tʃɔɪ/ and /ʃo/–/tʃo/, /ʊʃ/–/ʊtʃ/ and /ɛʃ/–/ɛtʃ/, and /ʊk/–/ʊt/ and /ɔk/–/ɔt/. In all cases, one of the endpoints would be a real word (*show*, *etch*, and *ought*), and thus we would expect a lexical bias in these series. We would expect that listeners would be biased toward identifying items as the real words (Ganong, 1980), precluding any measurement of coarticulatory effects. Moreover, there would still be neighborhood biases in these series, in addition to these lexical ones: /ʃo/ and /ʃɔɪ/ have larger neighborhoods than do /tʃo/ and /tʃɔɪ/; /ʊtʃ/ and, especially, /ɛʃ/ have larger neighborhoods than do /ʊʃ/ and /ɛʃ/; and both /ɔt/ and /ʊt/ have larger neighborhoods than do /ʊk/ and /ɔk/. Thus, eliminating the critical consonant from the series above would not allow us to test for coarticulatory effects in the absence of lexical and neighborhood effects.

GENERAL DISCUSSION

First, to summarize our results, during nonword (and word) recognition, all words similar to the target were activated. When lexical neighborhoods were computed according to the one-phoneme change rule described by Luce (1987), the lexical decision and phoneme identification tasks showed consistent influences of variation in lexical neighborhoods on listener responses. This influence of neighborhood was present even though the size of the cohort for the items was controlled. In Experiments 3 and 4, all possible influences of cohort were eliminated by using nonwords whose only neighbors differed in the initial phoneme. Despite the lack of any difference in cohorts, an influence of lexical neighborhood on phoneme perception was found. In contrast, we found no evidence for any influence of lexical cohort (words with the same

initial sound) in the two lexical decision experiments. Most notably, in Experiment 2, when overall lexical neighborhood size was controlled and cohort size was manipulated, cohort size did not influence the listeners' lexical decision responses to nonwords.

These results extend previous findings regarding the effects of lexical neighborhoods by demonstrating that when neighborhood effects occur, they include influences of neighbors that do not match word initially. That is, even when a word mismatches a target on its first phoneme, it still can be partially activated during word recognition. Although this might be expected in fluent speech, where there are no pauses between words, the results of Experiments 1–4 show that this is also the case for nonwords presented in isolation where the word or nonword onset was not ambiguous. This contradicts some previous findings, such as those of Marslen-Wilson and Zwitserlood (1989), who used a cross-modal priming paradigm and found that words that mismatched in their initial phoneme did not show significant activation. However, they also reported that rhyme primes might have detectable priming effects in some situations and, in fact, demonstrated a nonsignificant effect in the appropriate direction in their own results. One possible conclusion is that neighbors that mismatch at onset are activated, but to a lesser degree than those without the mismatch (such as the comparison items in their study). This lesser degree of activation might prevent individual neighbors from showing significant priming effects but would not prevent them from combining (or "ganging") to influence perception as a group.

In contrast, Connine et al. (1993) found no difference in the size of a priming effect for items mismatching at onset versus later in a word. Moreover, the results from Experiment 2 also suggest that items matching word initially do not have any effect over and above that of items that match elsewhere. Perhaps items that mismatch initially are activated, but this activation takes longer to build up than does activation for items that match initially. In this case, the precise timing relations between when a target item is presented and when responses occur might be a critical factor in terms of whether such effects are found in individual experiments.

The influence of lexical neighborhoods that mismatched on the first phoneme was seen in both lexical decision (Experiments 1 and 2) and phoneme identification (Experiments 3 and 4) results, so the effect is not a result of idiosyncrasies of the task, choice of stimuli, or representation used by listeners to make their decision. Finally, all of the neighbors in the series tested in Experiments 3 and 4 mismatched the targets in the initial phoneme, and this mismatch was in two or more features. This rules out models of lexical access such as the original version of cohort, in which only lexical candidates that matched the target on the initial phoneme would be activated. It also rules out any model in which a cohort is activated with only a minimal (one feature) initial mismatch being allowed.

The present results suggest that the onset of words may not play a unique role in word recognition (see also Nooteboom & van der Vlugt, 1988, for further support of this idea). Not only did neighbors that mismatched at word onset influence perception, but the results from Experiment 2 suggest that items that match at onset do not have any overall greater influence than those that mismatch at onset. There may still be differential time courses for the activation of words that match at the onset of an item versus words that mismatch at onset (but match elsewhere). We might expect that words will become activated only when they match some critical proportion of the acoustic information coming into the system. A word that mismatches late (matches at onset) would reach that critical point sooner than one that mismatches early. This means that cohort words might show signs of activation at an earlier point in time. It is possible that subtle time course effects exist and that these may result in word onsets having a greater influence on word recognition than do word offsets in some or even many situations. However, word onsets do not appear to be processed in a separate, unique manner from word offsets. This proposal and our results are consistent with models of word recognition such as NAM (Luce & Pisoni, 1998) and Merge (Norris, McQueen, & Cutter, 2000).

There may be some evidence regarding this differential time course of activation in a comparison of the data for our *zook–zawt* series in Experiment 4 with the data in McQueen (1991) that show a lexical effect for a word- (syllable-) final contrast. With the *zook–zawt* series, an effect of lexical neighborhood was found in both listeners' fast and intermediate speed responses. McQueen, in his lexical effect study in which *fish–kish* and *fiss–kiss* series were used, found an influence of the lexical status of the syllable on his listeners' fast responses, but not on their intermediate (or slow) responses. At first glance, this result is a bit puzzling. As a general rule, neighborhood influences on phoneme perception have a more rapid time course and show up earlier in processing than do effects of lexical status. Newman et al. (2000) showed this for series with a syllable-initial contrast. On the basis of this, the influence of neighborhood with our syllable-final contrast should occur for the same RT partition as McQueen's lexical effect (the fast partition) or even earlier in processing. Clearly, the finding of a neighborhood effect for the intermediate speed partition does not match this expectation.

It is certainly the case that this comparison must be made with caution, since the phonetic contrast that listeners are identifying and the overall syllable durations differ between these studies. However, it is entirely plausible that the difference in the time course of the effects of neighborhood and lexical status reflects an influence of lexical cohort. In McQueen's (1991) study, the initial phoneme and the following vowel in each series were a part of the word for one endpoint of that series. Thus, as listeners heard /fɪ/ or /kɪ/, the neighborhood activation would include the words *fish* and *kiss* and their cohort. When the ambiguous final phoneme that could be identified as /s/ or as /ʃ/ had been processed, the lexical entries for the word ends and their cohorts had already been

partly activated. As a consequence, this partial activation caused the influence of lexical status (and cohort) to occur in the rapid responses that listeners made to the final consonant.

By comparison, in Experiment 4, the lexical neighbors for the series stimuli did not match the initial /z/ phoneme of the target. Consequently, the activation of the neighbors that ultimately influenced the listeners' responses to the final consonant began with the processing of the vowel and continued with the processing of the final consonant. When compared with the words and syllables used by McQueen (1991), the neighbors of the syllables of Experiment 4 built up lexical activation more slowly. As a consequence, the influence of neighborhood, which normally has a faster time course than the influence of lexical status, was delayed. The net result was that the influence of neighborhood for the *zook–zawt* stimuli was found for both the fast and the intermediate speed responses. A similar analysis could be made with the results of Experiment 3, where the pattern in the data, although weaker, was the same as that in Experiment 4.

This comparison of the lexical status results with our results is post hoc. However, the explanation is consistent with our interpretation of what constitutes a lexical neighborhood and the role of a cohort in word recognition. The influence of a cohort is confined to the time course of partial activation of words in the mental lexicon. Entries that match the target at onset start activation sooner than do neighbors that do not share the initial phoneme with the target. However, all neighbors are activated and contribute to the process of word and nonword recognition.

In summary, we have found consistent evidence that lexical neighborhoods influence the perception of nonwords and that words that are similar to any portion of a target item seem to contribute to lexical neighborhood effects. Finally, these results suggest that multiple activation is not merely an issue of lexical confusion prior to the arrival of a word's uniqueness point. Were this the case, we would expect that as soon as a lexical item was found to differ from the target, its activation level would drop, with the end result that items mismatching word initially would not show significant activation or influence perception. Instead, our results favor a view of word recognition that is consistent with the principle of least commitment, as described by Marr (1982) for vision. Massive activation of all similar words may be a way of compensating for potential mishearings. Activating items that are *similar* to the target items means that reinterpretations are available in case the favored or first interpretation of a sound turns out to be incorrect.

REFERENCES

ALLOPENNA, P. D., MAGNUSON, J. S., & TANENHAUS, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language*, **38**, 419-439.

CLUFF, M. S., & LUCE, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 551-563.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

COLE, R. A., JAKIMIK, J., & COOPER, W. E. (1978). Perceptibility of phonetic features in fluent speech. *Journal of the Acoustical Society of America*, **64**, 44-56.

COLE, R. A., & PERFETTI, L. A. (1980). Listening for mispronunciations in a children's story: The use of context by children and adults. *Journal of Verbal Learning & Verbal Behavior*, **19**, 297-315.

CONNINE, C. M., BLASKO, D. G., & TITONE, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory & Language*, **32**, 193-210.

ENGEN, T. (1971). Psychophysics I: Discrimination and detection. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's experimental psychology* (pp. 11-46). New York: Holt, Rinehart & Winston.

FOX, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 526-540.

GANONG, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 110-125.

GOLDINGER, S. D., LUCE, P. A., & PISONI, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory & Language*, **28**, 501-518.

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

LUCE, P. A. (1987). Neighborhoods of words in the mental lexicon. *Dissertation Abstracts International*, **47**(12-B, Pt. 1), 5078.

LUCE, P. A., & PISONI, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, **19**, 1-36.

LUCE, P. A., PISONI, D. B., & GOLDINGER, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122-147). Cambridge, MA: MIT Press.

MARR, D. (1982). *Vision*. San Francisco: Freeman.

MARSLEN-WILSON, W. D. (1987). Functional parallelism in spoken-word recognition. *Cognition*, **25**, 71-102.

MARSLEN-WILSON, W. D. (1990). Activation, competition and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.

MARSLEN-WILSON, W. D., MOSS, H. E., & VAN HALEN, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1376-1392.

MARSLEN-WILSON, W. D., & WELSH, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.

MARSLEN-WILSON, W. D., & ZWITSERLOOD, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 576-585.

MCQUEEN, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 433-443.

MILLER, J. L., & DEXTER, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 369-378.

NEWMAN, R. S., SAWUSCH, J. R., & LUCE, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 873-889.

NEWMAN, R. S., SAWUSCH, J. R., LUCE, P. A., & AUBIN, A. (2000). The relative time course of neighborhood and lexical effects [Abstract]. *Journal of the Acoustical Society of America*, **107**, 2857.

NOOTEBOOM, S. G., & VAN DER VLUGT, M. J. (1988). A search for a word-beginning superiority effect. *Journal of the Acoustical Society of America*, **84**, 2018-2032.

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

NUSBAUM, H. C., PISONI, D. B., & DAVIS, C. K. (1984). *Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words*

(Research on Speech Perception Progress, Rep. 10). Bloomington: Indiana University.

PITT, M. A., & SAMUEL, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 699-725.

RAAIJMAKERS, J. G. W. (2003). A further look at the "language-as-fixed-effect fallacy." *Canadian Journal of Experimental Psychology*, **57**, 141-151.

RAAIJMAKERS, J. G. W., SCHRIJNEMAKERS, J. M. C., & GREMMEN, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory & Language*, **41**, 416-426.

SAMUEL, A. (1986). Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology*, **18**, 452-499.

VAN DONSELAAR, W. (1996). Mispronunciation detection. *Language & Cognitive Processes*, **11**, 621-628.

VITEVITCH, M. S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 270-278.

VITEVITCH, M. S., & LUCE, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory & Language*, **40**, 374-408.

WALLEY, A. C. (1987). Young children's detections of word-initial and -final mispronunciations in constrained and unconstrained contexts. *Cognitive Development*, **2**, 145-167.

WESTBURY, C., BUCHANAN, L., & BROWN, N. R. (2002). Sounds of the neighborhood: False memories and the structure of the phonological lexicon. *Journal of Memory & Language*, **46**, 622-651.

NOTES

1. For each word in a neighborhood, we first multiplied the raw frequency (from Kučera & Francis, 1967) of each item by 10 and then took the logarithm (base 10) of that number. The multiplication was done so that items with a frequency of 1 would not end up with a log of 0 and, thus, no influence on the weighted neighborhood frequency. If the frequency of a word in the on-line dictionary was 0, it was replaced with a 1 before the computation.

2. Category boundary responses gave similar results. There was an overall shift in the category boundary [$t(19) = 3.45, p < .005$; $\eta^2 =$ .385]. After partitioning, there was a significant effect of both partition [$F(2,38) = 5.33, p < .01$; $\eta^2 = .025$] and series [$F(1,19) = 11.87, p < .005$; $\eta^2 = .105$], but no interaction ($F < 1$; $\eta^2 = .001$). This suggests that there was a consistent effect of neighborhood across all three speed ranges, as well as a trend toward an earlier *sh–ch* category boundary at longer RTs. Planned comparisons showed shifts in category boundaries in all three response categories [for fast responses $t(19) = 2.51, p < .05$; for intermediate speed responses $t(19) = 3.39, p < .005$; and for slow responses $t(19) = 1.90, p < .05$; η^2 values were .249, .377, and .160, respectively].

3. There was no effect of neighborhood on the category boundaries [$t(21) = 0.11, p > .20$; $\eta^2 = .001$], despite the significant effect by percentage of responses. The reason for this difference across methods of analysis is unclear. It may be that percentage of response measures are a more sensitive measure for this type of analysis, as has been claimed previously (Samuel, 1986). Or it may be that this is an indication that the effect is located only at the ends of the continuum, suggesting that it is more a bias than a perceptual change. The two-way ANOVA on category boundary data showed a significant effect of RT partition [$F(2,42) = 11.78, p < .0001$; $\eta^2 = .06$], so that there was a later category boundary at short RTs, but only a nonsignificant trend toward a neighborhood effect [$F(1,21) = 1.68, p > .10$; $\eta^2 = .004$] and no interaction ($F < 1$; $\eta^2 = .002$). The planned comparisons showed no effects of neighborhood in the category boundary data for any of the RT partitions [for fast responses, $t(21) = 1.09, p > .10$; $\eta^2 = .054$; for intermediate responses, $t(21) = 0.03, p > .10$; $\eta^2 < .001$; and for slow responses, $t(21) = 0.72$, $p > .10$; $\eta^2 = .024$].

4. A button on our button box broke in the fourth block of one of our participants; the data from the first three blocks were kept. Thus, this participant only gave nine responses to each stimulus.

5. The category boundary data looked identical to those for percentage of *k* responding. Overall, there was a significant shift in the category boundary [$t(23) = 5.14, p < .0001$; $\eta^2 = .535$]; the ANOVA showed a significant effect of neighborhood [$F(1,23) = 20.84, p < .0001$; $\eta^2 = .074$] and a significant interaction between neighborhood and partition [$F(2,46) = 13.08, p < .0001$; $\eta^2 = .070$], but no main effect of partition [$F(2,46) = 1.42, p > .10$; $\eta^2 = .009$]. The paired comparisons showed effects for the fast [$t(23) = 4.89, p < .0001$] and the intermediate [$t(23) = 3.79, p < .0005$] responses, but not for the slow responses [$t(23) = -1.00, p > .20$; $\eta^2 = .510$, .384, and .042, respectively].

APPENDIX A
Items From Experiment 1 and Their Onset-Match
and Full-Neighborhood Frequency-Weighted
Neighborhood Values

Onset-Match Set			Full-Neighborhood Set		
/væl/	11.72	16.45	/θæl/	0.00	4.73
/θɪf/	11.22	15.56	/vɪf/	1.30	5.64
/ðɛb/	13.12	14.90	/ðɔk/	0.00	11.81
/dʒɔdʒ/	5.09	5.09	/dʒɛdʒ/	7.95	15.53
/θuk/	2.83	4.87	/θub/	0.00	2.49
/dʒutʃ/	8.42	8.42	/θæŋ/	7.08	22.37
/θɚŋ/	10.12	10.12	/dʒɚtʃ/	2.78	11.62
/fæz/	13.67	25.92	/fuz/	12.77	31.32
/wæl/	15.98	20.71	/wæŋ/	3.26	19.55
/joŋ/	2.59	2.59	/jol/	3.56	26.98
/dʒɑl/	13.50	15.51	/jɑb/	1.60	16.16
/jɔb/	1.30	1.30	/θɔl/	6.49	35.17
/θɔtʃ/	19.47	9.40	/dʒɔtʃ/	2.20	5.11
/wut/	12.07	27.23	/vut/	3.88	19.03
/vɑl/	6.65	8.66	/wɑb/	5.56	21.11
/tʃɛb/	7.24	7.24	/tʃɛl/	10.39	34.31
/vɛb/	2.60	4.38	/vɛm/	1.00	8.46
/dʒɑm/	16.02	25.06	/vɑb/	1.60	16.16
/vɑum/	1.42	2.85	/dʒɑutʃ/	0.00	4.38
/jutʃ/	11.20	16.20	/jum/	12.20	27.37
Total	186.26	237.46	Total	83.62	339.31

Note—The onset-match set has a higher neighborhood value if only those items that match at word onset are considered neighbors. The full-neighborhood set has a higher neighborhood value than the onset-match set if all words that differ by a single phoneme are considered neighbors. Sets match for phonemes in each position.

APPENDIX B
Items From Experiment 2 and Their Onset-Match
and Full-Neighborhood Frequency-Weighted
Neighborhood Values

Onset-Match Condition					
High Onset-Match Set			Low Onset-Match Set		
/wæb/	2.78	11.34	/væb/	4.11	12.66
/vɑl/	6.65	8.66	/wɑb/	5.56	21.11
/tʃɛb/	7.24	7.24	/tʃɛl/	10.39	34.31
/ʃʌm/	9.29	32.19	/ʃʌdʒ/	5.96	11.15
/vɑum/	1.42	2.85	/vɛm/	1.00	8.46
/nutʃ/	16.75	16.75	/num/	18.56	32.73
/ðɛb/	13.12	14.90	/ðɔk/	0.00	11.81
/dʒɔdʒ/	5.09	5.09	/dʒɑutʃ/	0.00	4.38
/θuk/	2.83	4.87	/θub/	0.00	2.49
/dʒɛp/	5.06	6.06	/dʒɑɪp/	0.00	9.75
/vɑɪp/	9.61	19.35	/vɛp/	1.00	2.00
/lɛb/	19.14	20.92	/lɛl/	16.49	40.41
/zæl/	0.00	4.73	/zæb/	0.00	8.56
/dʒɑm/	16.02	25.06	/tʃɑm/	5.89	14.94
/tʃɪv/	11.19	18.03	/dʒɪv/	6.85	13.69
/dʒutʃ/	8.42	8.42	/dʒɚtʃ/	2.78	11.62
/θɚz/	5.60	7.81	/θuz/	0.00	22.37
/wɔɪl/	14.99	24.54	/nɔɪl/	8.16	17.72
/mɑut/	25.89	37.22	/vɑut/	9.59	20.93
/nʌz/	19.59	26.87	/mʌz/	15.26	22.53
/vɛb/	2.60	4.38	/θɛb/	0.00	1.78
/vɛf/	1.00	7.88	/zɛf/	0.00	6.88
/zɪf/	2.90	7.25	/vɪf/	1.30	5.64
/θɪp/	9.31	28.40	/vɪp/	1.30	20.38
/fɪf/	20.14	24.48	/ʃɪf/	13.47	25.07
/ʃop/	21.49	36.48	/fop/	14.77	29.72
/hædʒ/	31.06	32.75	/sædʒ/	19.39	21.09
/sʌv/	23.25	27.92	/hʌv/	22.02	26.68
Total	312.43	472.41	Total	183.84	460.87

Full-Neighborhood Condition					
High Neighborhood Set			Low Neighborhood Set		
/vɛm/	1.00	8.46	/dʒɑm/	16.02	25.06
/vɑb/	1.60	16.16	/vɑum/	1.42	2.85
/dʒɑutʃ/	0.00	4.38	/jutʃ/	11.20	11.20
/jum/	12.20	27.37	/θɛb/	0.00	1.78
/zɪf/	2.90	7.25	/vɪf/	1.30	5.64
/vɛf/	1.00	7.88	/zɛf/	0.00	6.88
/θɪp/	9.31	28.40	/vɪp/	1.30	20.38
/lɛl/	16.49	40.41	/lɛb/	19.14	20.92
/zæb/	0.00	8.56	/zæl/	0.00	4.73
/vut/	3.88	19.03	/wut/	12.07	27.22
/wɑb/	5.56	21.11	/vɑl/	6.65	8.66
/tʃɛl/	10.39	34.31	/tʃɛb/	7.24	7.24
/wɔɪn/	10.13	15.94	/hɔɪn/	7.03	8.16
/hus/	20.42	30.27	/wus/	5.18	15.02
/θuz/	0.00	22.37	/θɚz/	5.60	7.81
/dʒɚtʃ/	2.78	11.62	/dʒutʃ/	8.42	8.42
/jɑb/	1.60	16.16	/dʒɑl/	13.50	15.51
/nɔl/	10.36	35.37	/jɔb/	1.30	1.30
/dʒut/	3.46	10.95	/nut/	13.43	28.92
/pɪm/	19.35	32.05	/mɪm/	16.02	28.68
/mɛm/	17.19	24.61	/pɛm/	13.58	21.02
/nɑp/	16.08	30.52	/gɑp/	9.42	23.86
/gɛm/	11.86	19.32	/nɛm/	10.52	17.97
/zɛm/	1.00	8.46	/zɛp/	1.00	2.00
/zɑp/	1.00	15.44	/zɑm/	1.00	10.05
Total	179.55	496.28	Total	182.34	331.27

Note—Items are arranged so that those on the same line have maximal phonetic overlap.