Cognition, 52 (1994) 163–187

0010-0277/94/$07.00

The effect of subphonetic differences on lexical access

Jean E. Andruski[a], Sheila E. Blumstein[a,*], Martha Burton[b]

[a]*Department of Cognitive and Linguistic Sciences, Brown University, Providence, RI 02912, USA*
[b]*Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA*

Received February 2, 1993, final version accepted November 15, 1993

Abstract

This study investigated whether lexical access is affected by inherent acoustic variations that contribute to the identity of a phonetic feature and ultimately a phonetic segment. Two experiments were conducted to determine whether the magnitude of semantic (associative) priming in a lexical decision task is influenced when the acoustic manifestation of the initial voiceless stop consonant (specifically, voice onset time) of a prime word semantically related to a lexical decision target was systematically manipulated (e.g., prime: "king"; target: "queen"). Results showed no effect of the acoustic manipulations at the longer interstimulus interval (250 ms); however, at the shorter interstimulus interval (50 ms), the magnitude of semantic facilitation decreased as a function of the voice onset time manipulations. In addition, there was a tendency toward slower lexical decision latencies when the prime word had a real word voiced counterpart than when it did not. These results suggest that activation levels of words in the lexicon are graded, depending on the subphonetic shape of the input word. Results also suggest that words which are phonologically similar to the intended word candidate are activated to some extent, whether the input provides a relatively poor phonetic representation of the intended word or a good one.

*Corresponding author.
This work was supported in part by NIH Grant DC00142 to Brown University and by SSHRC (Social Sciences and Humanities Research Council of Canada) Grant 753-91-0012 to the first author. We would also like to thank three anonymous reviewers for their helpful comments.

SSDI 0010-0277(94)00600-C

1. Introduction

All theories of speech perception must ultimately relate to the language processing system, and as a result it is important to understand the extent to which the acoustic properties of speech interface with and inform higher levels of language processing. It is generally assumed that in spoken word recognition a pattern is generated from the acoustic-phonetic information in the speech signal and matched with a stored pattern or representation in the mental lexicon. Once pattern matching is accomplished, abstract knowledge about the word, including semantic information, may be accessed. Word recognition includes those processes which involve matching acoustic patterns in the waveform to stored patterns in the lexicon, while lexical access includes processes whereby information about a word, such as its meaning, is retrieved. Although it is generally assumed that a word in the lexicon is represented in terms of an abstract phonetic or spectral representation, it is not clear whether acoustic-phonetic detail in the speech waveform is retained through the process of word recognition and lexical access, or discarded prior to lexical access.

It is well known that the acoustic manifestations of speech are affected by many sources of variability. Some of the sources of variability reflect rule-governed, context-dependent effects which influence the identification of phonetic segments. For example, in English the duration of a vowel provides a cue to the voicing properties of a following stop consonant. Similarly, the occurrence of a nasalized vowel serves as a cue to the occurrence of a following nasal consonant. Other sources of variability reflect subphonetic variations known to occur in speech production; for example, the acoustic manifestations of voiceless stop consonants may vary along a range of voice onset time (VOT) values from about 40 to 150 ms, and the duration of a vowel serving as a cue to postvocalic stop voicing ranges from about 100 to 250 ms (Denes, 1955; Raphael, 1972). These latter variations are "random", but must fall within a range of parameter values that define the particular phonetic dimension or phonetic category.

Studies in speech perception have shown that listeners are, indeed, sensitive to these sources of variability, and that they influence the identification of the sound segments of language. In fact, it has been shown that low-level acoustic differences influence speech processing even when subjects judge that the phonetic identities of the sound segments are the same. For example, Pisoni and Tash (1974) showed that CV syllables which were identified the same phonetically were nonetheless treated differently by listeners. In particular, it took subjects longer to judge CV syllables "the same" when the VOT values of the initial stop consonant were different than when they were acoustically identical. Additionally, "different" judgements for syllables from different phonetic categories were faster when the degree of acoustic difference between the syllables was increased.

It is possible, however, that low-level acoustic differences of this type affect speech processing, including the identification and discrimination of phonetic features, segments or syllables as well as the time it takes to make such decisions, without affecting higher processes such as word recognition. For example, low-level variations in the speech signal might affect the rate at which features or segments are extracted, but have no influence on higher levels of processing once featural, segmental and/or syllabic decisions have been made. If access to the lexicon takes place via features and segments which have been abstracted from the acoustic waveform, then fine acoustic detail could be discarded as soon as the appropriate featural or segmental information has been determined, and the information has been passed on to the next level or "module". In such a model, there would be no effect of the acoustic fine structure on word recognition or lexical access. As Klatt (1979) points out, however, if a speech recognition system commits itself to segmental (or other low-level) decisions at an early stage of analysis, it will be difficult for other components of the system to recover from errors when they occur. Klatt suggests that "if possible, [a speech recognition system] should not make phonetic decisions prior to lexical hypothesis formation" (1979, p. 283). If human speech perception incorporates this principle, low-level acoustic detail could be expected to affect word recognition and possibly also access to abstract knowledge about the word.

Some evidence is available which suggests that fine acoustic details do affect the process of word recognition. Streeter and Nigro (1979) showed that when word-medial CVC sequences were altered by either removing the VC transition or substituting it with a transition that conflicted with the preceding CV transition, lexical decision times for the altered words were significantly longer than for the original, unaltered words. Since decision times for nonwords were unaffected by the manipulations, they suggested that the increased reaction times could not be attributed to differences in initial processing. Rather, the delay must come during lexical lookup. As they point out, this would support a process such as that proposed by Klatt, in which a whole-word representation, including its acoustic detail, is used to search the lexicon. Similar conclusions were drawn by Whalen (1983), who also showed an effect of acoustic-phonetic manipulations on words, but not on nonwords.

Although it seems clear that lexical lookup for words was affected by Streeter and Nigro's acoustic manipulations, it remains unclear whether the increase in reaction times should be attributed only to differences in the amount of time required to match the incoming stimulus word to its internal form representation, or whether reaction time differences may reflect differential levels of activation for the lexical items themselves. That is, when the acoustic input provides a relatively poor exemplar of a lexical item it may simply take longer to find a match in the lexicon. However, differences in decision times could instead be introduced because the best-matched lexical item is activated at a relatively low level, compared with its activation for a good exemplar of the word.

In addition, since no simple analogy exists in natural speech for the kinds of manipulations made in Streeter and Nigro's experiment, namely the removal of coarticulatory cues and the juxtaposition of conflicting cues, the possibility exists that lexical lookup is not similarly affected by natural variations in the speech waveform. Further evidence that differences in acoustic fine structure can affect word recognition is provided by Warren and Marslen-Wilson (1987, 1988; cf. also Marslen-Wilson, 1989). In a series of experiments in which listeners were asked to identify monosyllabic words based on successively larger gated portions of the initial consonant(s) and vowel, they showed that listeners constrained their lexical choice using partial temporal and spectral cues to the voicing, place and manner of the final consonant present in the preceding vowel. They concluded that listeners continuously monitor the acoustic fine structure in the speech signal and make immediate use of partial spectral and durational cues provided by context, in the processes which lead to word recognition. Similar results were obtained by Lahiri and Marslen-Wilson (1991), who showed that listeners use rule-governed phonetic/phonological variability in the process of word recognition. They demonstrated that English listeners used nasality in a vowel to anticipate the presence of a following nasal consonant – a phonetically based strategy – whereas Bengali listeners used the nasality in a vowel as a cue to the underlying phonemic status of the vowel. In contrast to the English listeners, Bengali listeners waited until later in the speech stream to identify a nasal consonant, since vowel nasalization is not a cue in Bengali, as it is in English, to the nasality of the following consonantal segment.

These results show that listeners can harness important contextual variations, and use them to anticipate what patterns in the mental lexicon will match the incoming speech signal. So-called "partial information" can act as an anticipatory cue to the identity of upcoming sound segments. However, it again remains unclear whether lexical entries are differentially activated depending on the strength or "goodness" of the partial information that is present at the time of lexical lookup.

So far, then, it appears that a variety of types of acoustic information (phonologically predictable, as in Marslen-Wilson and his colleagues' work, and unpredictable as in Streeter and Nigro) can exert an effect on word recognition. However, there is no clear evidence of differential activation within the lexicon itself based on the degree of accord or conflict between the long-term lexical representation and the temporary, internal pattern with which it is matched. Similarly, it is unclear whether the inherent variation that contributes to the identity of a phonetic feature, and ultimately a phonetic segment, in the process of word recognition can also affect lexical access. In the case of inherent variation, there is a range of acoustic values that characterize a particular phonetic parameter. Unlike coarticulatory or contextual effects, this range of parameters is not predictable from context. It does, however, reflect normal variation that occurs in the instantiation of a phonetic feature, and it may also determine the extent to which a particular parameter value represents a "good exemplar" of a particular phonetic category. For example, the VOT values for initial [t] may vary from approximately 40 to 150 ms, and the occurrence of any one VOT value is not predictable from the phonetic context. Those voiceless stops with shorter VOT values are closer, along a VOT continuum, to voiced stops, which have VOT values no longer than about 15–20 ms. The question is whether this inherent variability is used by the listener during lexical access, and in particular whether acoustic variations in parameters such as VOT can result in graded activation of lexical entries.

To explore this question, two experiments were conducted to determine whether semantic (associative) priming in a lexical decision task is influenced by the acoustic manifestation of the initial consonant of a prime word semantically related to a lexical decision target (e.g. prime: "king"; target: "queen"). To this end, the VOT of the initial voiceless stop of a set of prime words (e.g., "king") was systematically shortened such that it was one-third or two-thirds of the VOT value of the original exemplar stimulus, and yet all three prime words (the original, $\frac{1}{3}$ VOT and $\frac{2}{3}$ VOT) were perceived as beginning with a voiceless stop. We investigated whether such manipulations would affect the magnitude of priming for semantically related targets. If low-level acoustic differences do not affect lexical access, then semantic priming should occur in all three conditions, and the magnitude of priming should be similar in each condition. If the acoustic fine structure does affect lexical access, however, then the magnitude of semantic priming should decrease as the VOT decreases.

In addition, because recent research has shown that semantic facilitation is affected by the phonological relationship of other words in the lexicon to the prime word (Connine, Blasko, & Titone, 1993; Luce, Pisoni, & Goldinger, 1990; Milberg, Blumstein, & Dworetzky, 1988) we selected prime words in which the VOT contrast either provided a word competitor (e.g., "pet/bet") or resulted in a nonword ("king/ging"). It might be expected that the presence of a word competitor along the VOT continuum would affect the semantic facilitation results. In particular, there may be a reduction in priming, presumably due to inhibition caused by lexical competition, when the lexicon contains a voiced competitor.

Finally, the experiments examine the time course of priming of semantic associates by using different interstimulus intervals (ISIs) (250 and 50 ms). It is possible that acoustic fine structure affects lexical access at an early stage of processing, but that the effects do not persist throughout the time course of activation of lexical candidates.

PRELIMINARY PHONEME CATEGORIZATION TASK

In order to determine whether phonetic factors affect lexical access, it was first necessary to create a set of stimuli for which subjects showed a perceptual sensitivity to the acoustic phonetic manipulations, while at the same time maintaining the percept of the intended phonetic category. To this end, we developed a preliminary phoneme categorization task to allow us to select a set of stimuli which were identified by subjects as beginning with a voiceless stop and for which the subjects perceived within-category VOT differences.

2. Method

2.1. Stimuli

A phonetically trained male speaker of Canadian English who had resided in the Providence area for approximately five years produced a set of 100 one-syllable CV(C)C words. Fifty of the words began with /p/, /t/ or /k/. Three repetitions of each word were recorded. Of the /p,t,k/ words, half remained words if the voiceless stop was replaced with its voiced counterpart (e.g. "pet→bet") and half became non-words if the voiceless stop replaced with its voiced counterpart "king→ging"). The stimuli were recorded in a sound-treated room using a Nagra 4.2 tape recorder and a Shure SM 81 microphone and then digitized onto a MicroVax computer using a 20 kHz sampling rate and a 9 kHz low-pass filter setting. An acoustic analysis of the VOT of the /p,t,k/ words showed that the VOT of at least one of the three repetitions available for each word fell within a range of 60–100 ms. In order to achieve a relatively uniform test set, the token of each set was chosen whose VOT value fell closest to the midpoint of this range, that is, 80 ms. The resulting mean and range VOT values for the tokens included in the preliminary phoneme categorization task are summarized in Table 1. In addition, 50 words beginning with /b/, /d/ or /g/ were recorded by the same speaker for use as distractors. From the three recorded tokens of each /b,d,g/ word, the token which was most clearly enunciated was digitized.

The set of 50 /p,t,k/ words constituted the unaltered voiceless stimuli for the preliminary phoneme categorization task. Two altered versions of each /p,t,k/ word were then created by reducing the VOT of each word. In order to construct the reduced stimuli, the full VOT of each unaltered word was measured from the

Table 1. *Mean and range VOT (ms) for 50 unaltered /p,t,k/ words used in the preliminary phoneme categorization task*

	/p/ ($n = 17$)	/t/ ($n = 16$)	/k/ ($n = 17$)
Mean	79.6	78.6	80.3
Range	61.7–95.1	64.2–97.4	61.2–91.7

beginning of the burst to vowel onset. The halfway point between the two measurement cursors was designated as the VOT midpoint. For the first set of altered stimuli, one-sixth of the full VOT was removed from either side of the VOT midpoint, producing a token in which the original VOT had been reduced by one-third. These were designated the $-\frac{1}{3}$ ("minus $\frac{1}{3}$") stimuli. For the second set of altered stimuli, one-third of the full VOT was removed from either side of the VOT midpoint, providing tokens in which the original VOT was reduced by two-thirds. These were designated the $-\frac{2}{3}$ ("minus $\frac{2}{3}$") stimuli. All $-\frac{1}{3}$ and $-\frac{2}{3}$ stimuli were checked for transients or distortion which might have been introduced by the alteration process. No unusual noise was noted and no tokens were eliminated by the screening.[1]

The full set of stimuli for the preliminary phoneme categorization task consisted of the three versions of each 50 /p,t,k/ words (unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$) plus a distractor set of 50 /b,d,g/ words consisting of 17 words beginning with /b/, 16 words beginning with /d/ and 17 words beginning with /g/. The distractor words were included so that one-half of the test stimuli began with voiceless stops and one-half began with voiced stops. For presentation, three separate blocks of these 100 words were created. A different version of each /p,t,k/ word (unaltered, $-\frac{1}{3}$ or $-\frac{2}{3}$) was randomly assigned to each of the blocks. The resulting blocks each contained an approximately equal number of unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ files. Thus, each /p,t,k/ word appeared a total of three times, once in each block, and the three different versions of a /p,t,k/ word each appeared once. Each of the 50 /b,d,g/ distractor words appeared once in each block, and therefore occurred a total of three times in the experiment.

2.2. Subjects

Twenty-one students at Brown university were paid to participate in the task. All were native English speakers with no known hearing disorders.

[1]The decision to reduce VOTs by one-third and two-thirds was based on earlier pilot work which showed that reducing VOTS by absolute values of 20 and 40 ms gave inconsistent phonetic categorization results. Since bilabials tend to have the shortest and velars the longest VOTs of English stops, the effect of deleting an absolute VOT value of 20 or 40 ms had differential effects as a function of place of articulation of the stop. For bilabials, deletion of 20 ms brought some tokens relatively close to the boundary between voiced and voiceless stops. In certain cases deletion of 40 ms resulted in perception of a voiced bilabial token. For velars, however, deletion of 20 ms was often imperceptible, and with numerous tokens deletion of 40 ms did not bring the token close to the voiced/voiceless boundary. As a consequence, we decided to vary each token in relation to its own value by using a percentage of the actual VOT.

2.3. Procedure

Subjects were told they would hear a series of words starting with one of the six consonant sounds /b,d,g/ or /p,t,k/. They were required to press a button labeled "b, d, g" if the token started with a /b,d/ or /g/ sound, and to press a button labeled "p, t, k" if it started with a /p,t/ or /k/ sound. They were instructed to respond as quickly and accurately as possible. A complete testing session lasted approximately 20 min.

3. Analysis of results

The test stimuli were first scored for correct categorization. Those /p,t,k/ stimuli of which each version was correctly categorized at least 90% of the time were selected out for possible inclusion in the test set for the lexical decision experiment. There were a total of 34 such words. From this group, a set of 12 /p/ words, 8 /t/ words and 6 /k/ words were selected such that within each stop category half the words had nonword voiced counterparts ("king→ging") and half had real-word voiced counterparts ("pet→bet"). The mean correct categorization rate for the initial stop in the 26 selected /p,t,k/ words was 96.4%.

To determine whether subjects could perceive the within-category manipulations of VOT, an analysis was conducted of the reaction times (RTs) for the correct categorization responses of the 26 selected words in the three VOT conditions. Fig. 1 shows the results. Only correct categorization responses within two standard deviations of each individual subject's mean RT were included. As

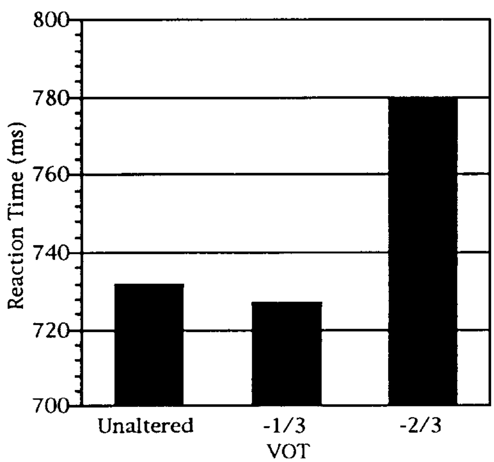

Figure 1. *Mean RTs to 26 /p,t,k/ words which were at least 90% correctly identified as containing a word-initial voiceless stop in the preliminary phoneme categorization task.*

the figure shows, RTs were fastest for the unaltered and $-\frac{1}{3}$ VOT conditions, and were considerably slower at the $-\frac{2}{3}$ VOT condition. A one-way repeated-measures analysis of variance (ANOVA) confirmed these findings, $F(2,40) = 21.88$, $p < .0001$. Post hoc Newman–Keuls tests showed that RTs to the $-\frac{2}{3}$ versions of the words were significantly slower than RTs to the $-\frac{1}{3}$ and unaltered versions ($p < .01$). RTs to the unaltered and $-\frac{1}{3}$ versions did not differ significantly ($p > .05$).

Thus, while subjects perceived both the unaltered and the manipulated VOT stimuli as voiceless over 90% of the time, their slower RTs in the $-\frac{2}{3}$ condition indicate that they are sensitive to the acoustic manipulations of VOT. The $-\frac{2}{3}$ VOT stimuli are presumably poorer exemplars of the voiceless phonetic category, and thus categorization responses are slower. Although VOT was reduced on average by more than 20 ms in the $-\frac{1}{3}$ condition, this manipulation was apparently not sufficiently great to affect the processing of these stimuli relative to the unaltered exemplars. The unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ versions of these 26 /p,t,k/ words chosen on the basis of the preliminary phoneme categorization task were used as test stimuli in Experiment 1.

EXPERIMENT 1

4. Method

4.1. Stimuli

Twenty-six real-word targets were preceded by four priming conditions to form pairs in which the first word was the prime and the second the target. These four conditions constituted the test items (see the Appendix for a complete list). In the first three test conditions the prime word was either the unaltered, $-\frac{1}{3}$ or $-\frac{2}{3}$ version of one of the 26 /p,t,k/ words selected in the preliminary phoneme categorization task. These words were semantically related to the following target word, for example, "$\text{peace}_{-1/3}$:war", "$\text{peace}_{-2/3}$:war", "$\text{peace}_{\text{unaltered}}$:war". For half of the prime words, the voiced counterpart was a non-word, for example, "peace → beace", and for half it was a real word, for example, "pet → bet". The fourth test condition consisted of the same set of target words paired with a set of semantically unrelated primes, for example, "hip:war". The semantically unrelated primes began with a sound other than /p/, /t/ or /k/. Targets were either one- or two-syllable words.

Four equivalent distractor conditions were also constructed. Targets in the distractor conditions were one- or two-syllable nonwords. The nonword targets were constructed from a new set of actual English words by replacing the initial sound or consonant cluster with some other sound or cluster which was acceptable according to English phonotactic constraints, for example, "faucet→daucet". These words were similar in phonological shape to the real-word targets. As with the test conditions, there were four distractor priming conditions: three contained a /p,t,k/ prime word, altered as described above to give a $-\frac{1}{3}$ and $-\frac{2}{3}$ set in addition to the unaltered set of distractor primes. All /p,t,k/ distractor primes were one-syllable CV(C)C words. As in the set of test primes, half the /p,t,k/ distractor primes were words if the initial voiceless stop was replaced with its voiced counterpart ("tip→dip") and half were nonwords ("tag→dag"). The fourth priming condition for the distractor set consisted of non-/p,t,k/ prime words paired with nonword targets, for example, "foot:chand".

The frequency of occurrence of the primes was controlled across the test and distractor sets. The word frequencies averaged 61 per million for the test set, and 62 per million for the distractor set (Francis & Kucera, 1982). In total, the experimental stimuli consisted of 208 trials, half of which were WORD responses and half of which were NONWORD responses. The stimuli were presented in four randomized blocks such that a different version of each prime word (unaltered, $-\frac{1}{3}$, $-\frac{2}{3}$, unrelated) was assigned to each one of the blocks. The ISI between prime and target words was set at 250 ms and the intertrial interval (ITI) was 3000 ms.

4.2. Procedure

Stimuli were presented to subjects at a comfortable listening level over Sony MDR-V2 or MDR-CD606 dynamic stereo headphones in a sound-treated room. Subjects were tested in groups of one, two or three. They were told they would hear pairs of words and nonwords, and were instructed to press a button labeled "Word" if the second item in the pair was an English word, or "Nonword" if it was not. They were asked to respond as quickly and accurately as possible. A practice set consisting of 5 stimulus pairs was presented before the 208-pair experiment set. Subjects were given no feedback on their responses to the practice tokens, but were given an opportunity to ask questions after the practice.

Following the lexical decision task, all subjects participated in a phoneme categorization task. This task was included to ensure that subjects perceived the initial consonant of the test prime stimuli as voiceless. In this task, subjects heard all three versions (unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$) of the 26 test prime words used in the lexical decision task. In addition, the phoneme categorization task contained 26 /b,d,g/ distractors selected from the preliminary phoneme categorization task. Each of the /b,d,g/ words was presented three times. Stimuli were randomized using the same method as described for the preliminary phoneme categorization task and subjects were given the same instructions as for that task.

A complete testing session including both the lexical decision task and the phoneme categorization task lasted approximately 45 min.

4.3. Subjects

Twenty-eight students at Brown University served as subjects and were paid for their participation. All were native speakers of English with no known hearing impairments.

5. Analysis of results

5.1. Phoneme categorization task

The data were analyzed for both correct categorization rate and mean RT to correct responses. Mean categorization rates and RTs for each version of the test words (unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$) are summarized in Table 2. Correct categorization rates were greater than 94% for all versions of the test words. As in the preliminary categorization task, the mean RT was considerably slower for the $-\frac{2}{3}$ versions of words than for the unaltered and $-\frac{1}{3}$ versions. Statistical tests replicated the results for the preliminary phoneme categorization task.

5.2. Lexical decision task

RT data from the lexical decision task were analyzed for mean latencies of correct lexical decisions to real word targets. The results are shown in Fig. 2. As the figure shows, subjects showed semantic facilitation in the lexical decision task; that is, they showed faster RTs to target words when the target was preceded by a

Table 2. *Mean correct categorization rates and RTs (ms) for the initial stop consonant of each version of the 26 selected test words*

VOT	% Correct	Mean RT
Unaltered	99.3	789
$-\frac{1}{3}$	98.5	786
$-\frac{2}{3}$	94.4	832

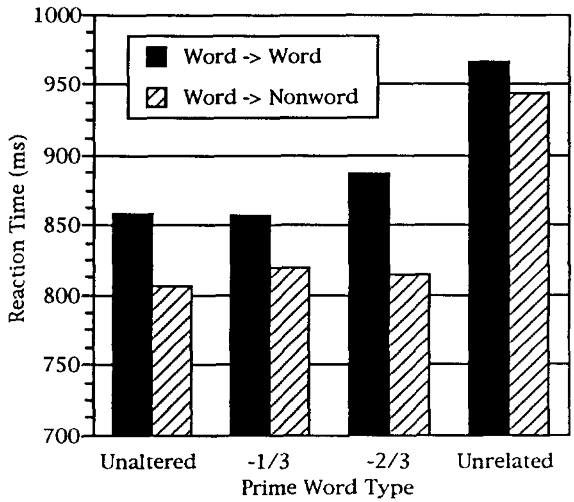

Figure 2. *Mean RTs to target words in the lexical decision task in Experiment 1. Black bars indicate targets preceded by primes whose voiced counterparts are real words ("pet→bet"). Shaded bars indicate targets preceded by primes with nonword voiced counterparts ("king→ging"). Unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ files are semantically related to the target.*

semantically related prime (the unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ conditions) than when it was preceded by a semantically unrelated prime (the unrelated condition).

Moreover, overall RTs were slower for targets preceded by primes which remained words if the initial stop was replaced with a voiced stop ("pet→bet") than for targets preceded by primes whose voiced counterpart was a nonword ("king→ging"). Within the group of semantically related primes (the unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ conditions) RTs look to be affected in the word→word condition as a function of the VOT manipulations. However, two-way repeated-measures subject and item ANOVAs with prime type (unaltered, $-\frac{1}{3}$, $-\frac{2}{3}$, unrelated) and prime word competitor (word→word, word→nonword) as within-subjects variables showed no significant interaction between the variables: subject $F(3, 81) = 1.93$, $p > .13$; item $F(3, 36) = .467$, $p > .7068$. A significant main effect for prime type was found in both ANOVAs: subject $F(3, 81) = 46.39$, $p < .0001$; item $F(3, 36) = 18.534$, $p < .0001$. Similarly, a significant main effect for prime word competitor was found in both ANOVAs: subject $F(1, 27) = 49.11$, $p < .0001$; item $F(1, 12) = 6.72$, $p < .0236$. The main effect for prime word competitor indicates that RTs to targets are significantly slower if primes remain words when a voiced stop is substituted for the initial voiceless stop ("pet→bet") than if they become nonwords. Newman–Keuls post hoc tests on the main effect for prime type in both ANOVAs showed that RTs to targets preceded by semantically related primes were significantly faster than RTs to targets preceded by semantically unrelated primes ($p < .01$). However, RTs were not significantly affected by VOT differences within the set of semantically related prime words ($p > .05$).

A further analysis was conducted on the RT data in order to determine whether the pattern of results that emerged was affected by the inclusion of several subjects who did not show perceptual sensitivity to the phonetic manipulation of VOT. Subjects who were not perceptually sensitive to differences between the unaltered and reduced VOTs on the phoneme categorization task would presumably fail to show any difference in the amount of facilitation that the unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ versions of the words produce on semantically related targets. A subset of 8 subjects who had slightly faster mean RTs to the $-\frac{2}{3}$ versions of words than to the unaltered versions on the phoneme categorization task were considered perceptually insensitive to the changes in VOT, and all data for these 8 subjects were eliminated from the second analysis. Results for the remaining 20 subjects are presented in Table 3. The pattern of results remains essentially unchanged. Both subject and item ANOVAs confirmed this finding. The main effects of prime type (unaltered, $-\frac{1}{3}$, $-\frac{2}{3}$, unrelated) and prime word competitor (word→word, word→nonword) were significant: subject F for prime type$(3, 57) = 37.451$, $p < .0001$; subject F for prime word competitor$(1, 19) = 29.942$, $p < .0001$; item F for prime type$(3, 36) = 16.36$, $p < .0001$; item F for prime word competitor$(1, 12) = 8.33$, $p < .0137$. There was again no significant interaction between lexical status and prime type: subject $F(3, 57) = 1.28$, $p > .29$; item $F(3, 36) = .503$, $p > .6824$.

Newman–Keuls post hoc tests on prime type indicated no significant differences in RTs due to reductions in the VOT of the prime word, but RTs to targets

Table 3. *Mean RTs (ms) to target words in the lexical decision task. The analysis in the last two columns excludes 8 subjects for whom the mean RT to the unaltered versions of words was longer than to the $-\frac{2}{3}$ versions of words on the phoneme categorization task*

Prime	All subjects included ($n = 28$)		Subjects who showed the phonetic effect ($n = 20$)	
	Word→word	Word→nonword	Word→word	Word→nonword
Unaltered	858	806	855	803
$-\frac{1}{3}$	857	820	846	821
$-\frac{2}{3}$	886	815	868	798
Unrelated	966	945	965	939

preceded by semantically related words were significantly faster than RTs to targets preceded by semantically unrelated words ($p < .01$).

6. Discussion

Although subjects identified the initial stop consonants in the test prime words as voiceless and showed perceptual sensitivity to the acoustic-phonetic manipulations as shown by the phoneme categorization task, the VOT manipulations did not affect the degree of semantic facilitation in the lexical decision task. Although semantic facilitation appeared to be affected by VOT manipulations in the word→word condition, the results were not significant. Nevertheless, the overall slower RTs in the word→word condition suggest that the presence in the lexicon of a voiced word candidate that rhymes with the prime increases overall RT without having an effect on the amount of semantic facilitation. These results suggest that such word candidates compete with the prime word, slowing lexical decision times. Taken together, the results of Experiment 1 suggest that low-level acoustic differences do not affect lexical access. If lexical access were affected by acoustic fine structure, then the magnitude of semantic priming should have decreased as the initial consonant of the prime word became a poorer exemplar of the phonetic category. It is possible, however, that acoustic fine structure does affect lexical access, but the effect is too short-lived to be observed at prime/target intervals as long as 250 ms. Experiment 2 explores this possibility.

In shortening the ISI for Experiment 2, it was necessary to choose an interval which would take into consideration limits on the duration of auditory memory, while at the same time avoiding the potential effects of backward masking of the target on the preceding prime. With regard to the duration of auditory memory, it is possible that any effects of acoustic fine structure on lexical access are limited by the brief period of time during which the fine structure is retained in a precategorical auditory store. Estimates of the maximum duration of the essentially unanalyzed versions of stimuli in an auditory store range from approximately 200 to 300 ms (Cowan, 1984; Massaro, 1972; Pisoni, 1973). Given the limited time during which acoustic detail is apparently available, an ISI of 250 ms may be too long to detect any effect of acoustic fine structure. In shortening the ISI, however, the possibility of the target word imposing backward masking on the prime must also be considered. Massaro (1970) showed that the backward masking effects of one pure tone on another decrease to asymptotic levels when stimulus *onsets* are separated by at least 250 ms. Similarly, the backward masking effects of white noise on an acoustic click are largely dissipated 50 ms after click offset, and reach asymptotic levels when the interval between white noise onset and click offset is about 250 ms (Wilson & Carhart, 1971). Thus, an ISI of 50 ms was considered sufficient to eliminate the possibility of the target word imposing backward masking on the prime.

EXPERIMENT 2

7. Method

7.1. Stimuli and procedure

The stimuli and procedures for the lexical decision task in this experiment were identical to those in Experiment 1 except that the ISI between prime and target was 50 ms. In addition to the lexical decision task, all subjects again participated in the phoneme categorization task. The test materials and procedures were the same as in Experiment 1 except that half of the subjects performed the phoneme categorization task prior to the lexical decision task and half performed it after the lexical decision task (in Experiment 1, all subjects performed the phoneme categorization task as a post-test).

7.2. Subjects

Twenty-eight students at Brown university were paid for their participation in Experiment 2. None had participated in Experiment 1. All were native speakers of English with no known hearing impairment. Subjects were again tested in groups of one, two or three and were given the same instructions as for Experiment 1.

8. Analysis of results

8.1. Phoneme categorization task

The data were analyzed for mean correct categorization rate and for mean RT to correct responses. The results are summarized in Table 4. All versions of the test words were correctly identified by at least 95% of the subjects. RTs to the $-\frac{2}{3}$ versions of /p,t,k/ words were again slower than RTs to the unaltered and $-\frac{1}{3}$ versions. RTs to the unaltered and $-\frac{1}{3}$ versions of words did not differ. Statistical tests replicate the results for both the preliminary phoneme categorization task and the categorization task for Experiment 1.

Table 4. *Mean correct categorization rates and RTs (ms) for 28 subjects on test words in the phoneme categorization task for Experiment 2*

VOT	% Correct	Mean RT
Unaltered	98.4	794
$-\frac{1}{3}$	98.8	783
$-\frac{2}{3}$	95.6	814

8.2. Lexical decision task

As in Experiment 1, data for the test items from the lexical decision task were analyzed for mean latencies of correct lexical decisions to real-word targets. The results are shown in Fig. 3. Subjects again showed semantic facilitation, as can be seen from the faster RTs to targets preceded by semantically related primes (i.e., the unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ primes) than to targets preceded by semantically unrelated primes. Within the set of semantically related primes, overall RTs were again slower for targets preceded by primes which remained words if the initial stop was replaced with a voiced stop ("pet→bet") than for targets preceded by primes whose voiced counterpart was a nonword ("king→ging"). Moreover RTs were somewhat slower in both of the prime word competitor conditions (word→word, word→nonword) when the VOT of the prime was reduced by two-thirds.

A repeated-measures ANOVA was conducted with prime type (unaltered, $-\frac{1}{3}$,

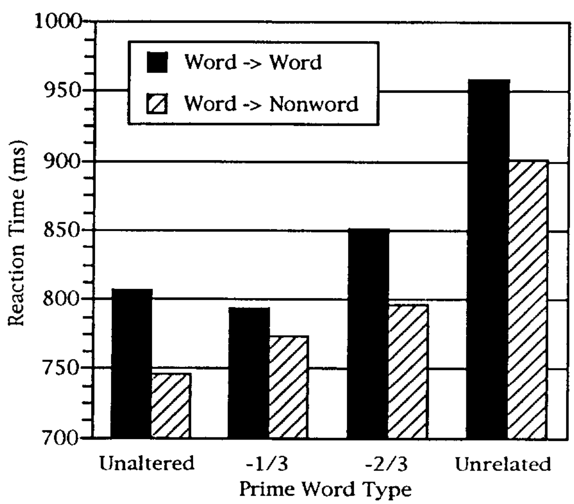

Figure 3. *RTs to target words on the lexical decision task for Experiment 2. Black bars indicate targets preceded by primes whose voiced counterparts are real words ("pet→bet"). Shaded bars indicate targets preceded by primes whose voiced counterparts are nonwords ("king→ging"). Unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$ files are semantically related to the target.*

$-\frac{2}{3}$, unrelated) and prime word competitor (word→word, word→nonword) as within-subjects variables, and ordering of the phoneme categorization task relative to the lexical decision task (pretest or post-test) as between-subjects variable. There were significant main effects for prime type, $F(3,78) = 6.65$, $p < .0005$, and prime word competitor, $F(1,26) = 5.84$, $p < .023$. No significant interactions were found. The main effect for pre/post-test was not significant: $F(1,26) = .006$, $p > .94$. Post hoc Newman–Keuls tests revealed that the significant effect for prime type was due to faster RTs to targets preceded by semantically related primes compared to unrelated primes ($p < .01$). In addition, RTs to targets preceded by $-\frac{2}{3}$ primes were significantly slower than RTs to targets preceded by unaltered and $-\frac{1}{3}$ primes ($p < .05$).

A two-way repeated-measures item ANOVA (Prime type × Prime word competitor) was also conducted. Significant priming effects were found, replicating the results of the subject analysis: $F(3,36) = 33.648$, $p < .0001$. Newman–Keuls tests showed that, as in the subject analysis, RTs to targets preceded by semantically related primes were faster than RTs to targets preceded by unrelated primes, and also that RTs to targets preceded by $-\frac{2}{3}$ primes were slower than RTs to targets preceded by unaltered and $-\frac{1}{3}$ primes. As in the subject analysis, the interaction between variables did not reach significance. Finally, the main effect for prime word competitor just failed to reach significance in the item analysis: $F(1,12) = 4.396$, $p < .0579$. This result suggests that subjects were generally slower in responding to targets preceded by primes with a real-word voiced counterpart ("pet→bet") than to targets preceded by primes with a nonword voiced counterpart ("king→ging").

These results suggest that the acoustic manipulations do have an effect on lexical access. However, it is also possible that the results reflect misperceptions of the initial consonants of the prime words as voiced rather than voiceless. That is, if a subject perceived the initial stop in "peace$_{-2/3}$" as a /b/, no semantic facilitation would necessarily be expected for the target word "war". Since RTs to targets preceded by semantically unrelated primes are slower than to targets preceded by related primes, it is possible that the slower overall RTs to the $-\frac{2}{3}$ primes result from misperception of some initial stops. Results indicate that 4.4% of the initial stops of the $-\frac{2}{3}$ tokens were misperceived as voiced stops in the phoneme categorization task. A Wilcoxon matched-pairs sign-ranks test indicates that this difference in error rates between unaltered and $-\frac{2}{3}$ versions of words is significant ($p < .01$). Thus, in order to reduce the possibility that the pattern of results was affected by misperceptions, each prime word whose initial stop was misidentified on the phoneme categorization task was eliminated from the lexical decision data for that particular subject. For example, if the initial stop of peace$_{-2/3}$ was misperceived as a /b/ on the phoneme categorization task, the prime-target pair "peace$_{-2/3}$:war" was eliminated from the subject's lexical decision data.

A comparison of RTs for the full data set and the data set with misperceived items removed showed that the largest change in mean RT for any stimulus condition was 3 ms. Statistical tests on the reduced data set confirmed that results were unchanged from the full data set.

Since it is clear from the results of the phoneme categorization task that subjects are slower at identifying the initial voiceless stop in the $-\frac{2}{3}$ primes than in the unaltered and $-\frac{1}{3}$ primes, another possible explanation for the results of the second lexical decision experiment remains. Specifically, the slower RTs to targets preceded by $-\frac{2}{3}$ primes may reflect slower processing times for the primes themselves, rather than indicating lesser amounts of semantic facilitation for the targets. Two different tests were conducted in order to assess this possibility. First, RTs to Nonword targets in the distractor set were examined. RTs to nonword targets preceded by $-\frac{2}{3}$ primes were compared to RTs to nonword targets preceded by unaltered primes. If the slower RTs to *test* targets preceded by $-\frac{2}{3}$ primes reflect increased processing requirements for the initial stop consonant, rather than reduced semantic facilitation, then RTs to nonword *distractor* targets preceded by $-\frac{2}{3}$ primes should be similarly affected. As shown in Table 5, the mean RT to nonword targets preceded by $-\frac{2}{3}$ primes was 15 ms slower than the mean RT to nonword targets preceded by unaltered primes. A paired, two-tailed t-test indicates that the 15 ms difference in means is not significant ($p > .14$).

In addition to examining the effect of prime word VOT on *nonword* targets, a second test examined the effect of the VOT manipulation on RTs to semantically unrelated *word* targets. If the slowed RTs to semantically related word targets in Experiment 2 simply reflect slower processing times for acoustically manipulated prime words, then slower RTs should also be obtained when a prime is followed by a semantically unrelated word. To explore this question, the full set of unaltered and $-\frac{2}{3}$ primes was recombined with the target words to provide a set of unrelated pairs of primes and real-word targets. An equal number of primes followed by nonword targets were also included, to give a new lexical decision task, again with half WORD and half NONWORD responses. Twenty-eight new subjects heard this set of stimuli. The results are presented in Table 6. No difference in RTs was found to word targets preceded by unrelated primes with unaltered or $-\frac{2}{3}$ VOTs. A similar lack of difference in RTs is shown for the nonword target responses. Both subject and item ANOVAs confirmed these

Table 5. *Mean and range RTs (ms) to nonword targets from the distractor set*

Prime word type	Mean RT	Range
Unaltered distractor	958	701–1290
$-\frac{2}{3}$ distractor	973	700–1291

Table 6. *Mean RTs (ms) to word targets preceded by semantically unrelated prime words, and to nonword distractor targets*

Prime word type	Word targets	Nonword targets
Unaltered	902	983
$-\frac{2}{3}$	903	981

results. There was no main effect for prime type (unaltered, $-\frac{2}{3}$), indicating that both nonwords and semantically unrelated words were unaffected by the acoustic manipulations made to prime word VOT: subject $F(1, 29) = .00235$, $p > .9962$; item $F(1, 25) = .165$, $p > .6877$. A significant main effect was found for target type (word, nonword), subject $F(1, 27) = 32.489$, $p < .0001$; item $F(1, 27) = 12.489$, $p < .0017$, indicating that nonword responses were slower than word responses. No significant interaction was found.

It would appear, then, that differences in RTs to the test target words in Experiment 2 are not due either to variations in the amount of time or resources required to process the preceding prime word, or to errors in perception of the initial voiceless stop of the prime. Instead, when the VOT of a prime word is reduced by two-thirds, the level of activation for that lexical item may be decreased in comparison with activation by an unaltered token. In turn, this may result in decreased activation levels for semantically related lexical items, leading to slower RTs for targets preceded by $-\frac{2}{3}$ primes on the lexical decision task.

9. Discussion

The results of Experiment 2 suggest that low-level acoustic differences can affect lexical access. However, these effects appear to be short-lived, since they emerged only when the ISI was shortened from 250 ms to 50 ms.

Phonetic effects emerged when the VOT of the initial stop consonant of the prime word was reduced by two-thirds. Importantly, significantly slower RTs were obtained for the full data set as well as for a reduced data set controlling for subjects' misperceptions of the initial voiceless stop of the prime word. Additional tests indicated that RTs to word targets preceded by semantically unrelated primes and RTs to nonword targets were unaffected by manipulation of prime word VOT. The lexical status of the prime word's voiced counterpart did affect subjects' overall RTs, but since item analysis results for this effect were not significant it must be interpreted with caution. Subjects showed slower RTs to targets when the voiced counterpart of the prime was a word than when it was a nonword. This suggests that the presence of a voiced rhyming counterpart in the lexicon can contribute to lexical inhibition of voiceless prime words.

Semantic facilitation was obtained in all three semantically related conditions (unaltered, $-\frac{1}{3}$ and $-\frac{2}{3}$), indicating that the prime word facilitates access to the target word in each of these conditions. The fact that significantly less semantic facilitation was obtained in the $-\frac{2}{3}$ condition compared with the unaltered and $-\frac{1}{3}$ conditions suggests that the acoustic fine structure of the prime word affects lexical access. In particular, the extent to which the VOT of the initial stop of the prime word is a good exemplar of the voiceless phonetic category appears to affect the level of activation of the prime word in the lexicon. That the results do not simply reflect a greater amount of time or additional resources required to process the "poor" exemplars of the prime stimuli is shown by the failure to obtain similar results for nonword targets preceded by /p,t,k/ primes which had also undergone the VOT manipulations, or to obtain these results for word targets preceded by semantically unrelated primes. Thus, the reduced semantic facilitation effects seem to relate to processes involved in lexical access, particularly the level of activation of the prime word.

GENERAL DISCUSSION

The results of Experiments 1 and 2 suggest that the representation which is first abstracted from the speech waveform incorporates the information contained in acoustic fine structure, including subphonetic variations; that this low-level acoustic information affects initial contact with the lexicon; and that activation levels in the lexicon are graded. The results also indicate that the effects of subphonetic variation on lexical access are brief, emerging by the methods used here at 50 ms and disappearing by 250 ms. Thus, the effects of acoustic fine structure on lexical access appear to diminish over the course of time.

It is worth noting that the effects of acoustic fine structure on semantic facilitation occurred whether the prime word had a real-word voiced counterpart ("pet/bet") or not ("king/ging"). In other words, at 50 ms, less semantic facilitation occurred with both "pet$_{-2/3}$" and "king$_{-2/3}$", compared to the unaltered exemplars of "pet" and "king". Similarly the effects of altered VOT on semantic facilitation disappeared for both types of stimuli at an ISI of 250 ms.

Importantly, however, the existence of a real-word voiced counterpart for the prime stimuli also had an effect on overall lexical decision latencies. Subjects showed slower decision latencies when there was a real-word counterpart for the prime than when there was not. This tendency toward slower latencies suggests that words which are phonologically similar to the intended word candidate are also activated to some extent during lexical access, whether the input provides a good phonetic representation of the intended word, as with the unaltered primes, or a poorer one, as with the $-\frac{2}{3}$ primes. Presumably, the presence or absence of a real-word counterpart contributes to the density or size of the set of activated lexical candidates. The results of the current experiments suggest that phonetic variants at word onset, and particularly the acoustic parameters for voicing, play a role in defining the set of competing lexical items.

Thus, there appear to be two independent factors contributing to the lexical decision effects obtained in Experiments 1 and 2. The first concerns the initial level of activation of a word candidate, and the second concerns the presence of other possible word candidates and their effect on overall lexical decision latencies. Our results indicate that relatively fine variations in VOT influence initial levels of activation in the lexicon. These results have important implications for current models of word recognition and lexical access. Though most models include graded activation at one or more levels leading to word recognition, none explicitly predict that such manipulations will produce graded activation up to and including the lexical level.

At first glance it may appear that in order to guarantee that low-level acoustic detail affects lexical activation levels, a model such as LAFS will be required. In the LAFS model of speech perception (Klatt, 1979, 1980, 1986, 1989), the speech waveform is transformed into a series of spectra and the spectra interact directly with the lexicon. However, it is possible that the results obtained here reflect, not the retention of actual acoustic detail, but rather the retention of information regarding the level of certainty of a particular decision. For example, if the amount of activation passed by two segments (say /p/ and /b/) to the lexicon is allowed to vary based on the likelihood that the speech waveform contains either of those segments, then in a minimal sense low-level acoustic information is retained up to the time of lexical access. In this case, strong competition between /p/ and /b/ to represent the same stretch of speech would imply that voicing is equivocal in the underlying waveform. No information remains about which or how many of the acoustic cues to voicing are indeterminate, but the indeterminacy of the underlying cues to voicing could nevertheless affect lexical activation levels. The information contained in acoustic detail is lost entirely only when a winner-takes-all decision is made, whether at the lexical level or at some intermediate level.

Cohort and TRACE are among the current models which include one or more intermediate levels of analysis, at which acoustic detail might be discarded. To our knowledge, there is currently no data, for these or any other models of lexical access, comparing how fine, sub-featural differences in the input are reflected at the lexical level (e.g., whether a given word node is differentially activated when the input contains two different phonetic exemplars of /p/).

TRACE (Elman, 1989; Elman & McClelland, 1986) involves three successive processing levels: the feature, phoneme and word. The level of activation of a given node indicates the degree to which the current input supports the conclusion that that feature, phoneme or word is present in the waveform. Since activation levels are allowed to vary continuously and more than one node may pass activation to a higher level, the presence of intermediate processing levels in TRACE does not oblige the system to discard the information available in low-level acoustic detail before lexical access. What needs to be explicitly accounted for, however, is whether feedback acts to "flatten" the (presumably small) differences in feature activation levels that are produced by subphonetic acoustic variations *before* the lexicon is reached, or whether these differences are at least briefly reflected by graded levels of activation in the node for the intended word.

With respect to the cohort model (Marslen-Wilson & Tyler, 1980), the crucial question is whether activation of features is continuous or binary on/off. In his 1987 revision of the cohort model, Marslen-Wilson suggests that input representations may consist of feature matrices. This allows predictable variations in phonetic structure to be used to access the lexicon. Hence in English, nasalization in a vowel can be used to anticipate the presence of a following nasal consonant. However, no role is specified for variations in sound structure which are not predicted by phonological rule, or which are smaller than a feature. If input representations consist of binary features, variations like those made to VOT in this experiment would be discarded at the input level, and could therefore not affect lexical activation.

Turning to the prime word competitor effect, our results showed a tendency toward slower RT latencies for targets that were preceded by primes which have a real-word voiced counterpart ("pet/bet") in comparison with primes which have a nonword voiced counterpart. In contrast with these findings, the cohort model predicts that, when the feature values of the initial input sound are unambiguous, only those lexical items which begin with the same segment as the input word will be activated (Marslen-Wood & Zwitserlood, 1989). Thus, cohort clearly suggests that an unaltered token that begins with /p/ in our experiment should activate only lexical items that begin with /p/. Since the initial segment of each $-\frac{2}{3}$ token was identified better than 90% correctly on each phoneme categorization task, it seems likely that the same prediction would be made for the set of $-\frac{2}{3}$ primes. However, our results indicate that in both cases the presence of a real-word voiced counterpart can affect lexical access. We suggest that this is because, when a rhyming word that begins with the voiced counterpart of the initial sound is present in the lexicon, it is also activated, regardless of the "goodness" of the initial voiceless stop, thus competing with the intended prime. This conclusion is also supported by the results of Milberg et al. (1988) and Connine et al. (1993), who found that nonwords which differed by one or two linguistic features from a base word significantly primed the semantic associates of the base word.

Models which do predict that voiced counterparts will contribute to a lexical competition effect include TRACE, the neighborhood activation model (Cluff & Luce, 1990; Luce et al., 1990) and the fuzzy choice model (Massaro, 1987, 1989; Massaro & Cohen, 1991). It remains unclear what acoustic, phonetic or phonological parameters contribute to lexical competition effects. However, these models point out the interesting possibility that the cohort or neighborhood of words that is activated may include any word which contains a segment that is auditorily confusable with the actual input, in any position. Further research is clearly needed to determine how factors such as phonological distance, word frequency, number of shared segments and features, and the position of segmental and featural differences within a word affect both membership in the competitive cohort and levels of competition.

A final point should be made regarding the origin of the differential facilitation effects found here. While the current studies were not designed to distinguish among the processing mechanisms for semantic facilitation, we favor spreading activation as the major source. There has been much discussion in the literature regarding the processing components contributing to semantic facilitation effects in lexical decision tasks, and there is evidence to suggest that semantic facilitation emerges as a consequence of both automatic and attentional processes.

The automatic, prelexical component, generally called spreading activation, assumes that words which have some semantic or associative relationship are connected in the lexicon. Activation of a given prime word spreads to its semantic associates, resulting in faster lexical decision times when those associated words are used as targets.

Several attentional mechanisms have also been proposed. On hearing the prime word, listeners may generate a set of potential target words, resulting in faster RTs when the actual target is a member of the set generated. It has also been proposed that, once a target word has been accessed in the lexicon, the retrospective awareness of a relationship between it and the prime may speed word decision times relative to nonword decision times (Neely, 1991; Neely & Keefe, 1989). Note that both of these attentional processes require the listener to make use of the lexically activated prime. Thus, if any of these mechanisms is the source of the differential semantic facilitation effects found here, our results indicate that the prime word and/or its associates must show graded activation at either the prelexical or postlexical processing stages, or both. According to current theory, however (Neely, 1991), the influence of automatic spreading activation can be expected to diminish across time, whereas the influence of attentional processes should increase as time passes. Since our results show a decrease in differential priming effects from the 50 ms ISI to the 250 ms ISI, they are most consistent with the predictions for spreading activation.

References

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 and Performance*, *16*, 551–563.

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 and Language*, *32*, 193–210.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, *96*, 341–370.

Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America*, *27*, 761–768.

Elman, J.L. (1989). Connectionist approaches to acoustic/phonetic processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 227–260). Cambridge, MA: MIT Press.

Elman, J.L., & McClelland, J.L. (1986). Exploiting lawful variability in the speech wave. In J.S. Perkell & D.H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 360–385). Hillsdale, NJ: Erlbaum.

Francis, W.N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Klatt, D.H. (1979) Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*, 279–312.

Klatt, D.H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R.A. Cole (Ed.), *Perception and production of fluent speech* (pp. 243–288). Hillsdale, NJ: Erlbaum.

Klatt, D.H. (1986). Problems of variability in speech recognition and in models of speech perception. In J.S. Perkell & D.H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 300–319). Hillsdale, NJ: Erlbaum.

Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169–226). Cambridge, MA: MIT Press.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*, 245–294.

Luce, P.A., Pisoni, D.B., & Goldinger, S.D. (1990). Similarity neighborhoods of spoken words, In G. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122–147). Cambridge, MA: MIT Press.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. In U.H. Frauenfelder & L.K. Tyler (Eds.), *Spoken word recognition* (a *Cognition* special issue), (pp. 71–102). Cambridge, MA: MIT Press.

Marslen-Wilson, W. (1989). Access and integration: projecting sound onto meaning. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 3–24). Cambridge, MA: MIT Press.

Marslen-Wilson, W., & Tyler, L.K. (1980). The temporal structure of spoken language understanding, *Cognition*, *8*, 1–71.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 576–585.

Massaro, D.W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, *85*, 411–417.

Massaro, D.W. (1972). Perceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, *79*, 124–145.

Massaro, D.W. (1987). *Speech perception by eye and ear: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

Massaro, D.W. (1989). Testing between the Trace model and the Fuzzy Logical Model of speech perception. *Cognitive Psychology*, *21*, 398–421.

Massaro, D.W., & Cohen, M.M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology*, *23*, 558–614.

Milberg, W., Blumstein, S., Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, *26*, 305–308.

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J.H., & Keefe, D.E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 207–248). New York: Academic Press.

Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, *13*, 253–260.

Pisoni, D.B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics*, *15*, 285–290.

Raphael, L.J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America*, *51*, 1296–1303.

Streeter, L.A., & Nigro, G.N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, *65*, 1533–1541.

Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics*, *41*, 262–275.

Warren, P., & Marslen-Wilson, W. (1988). Cues to lexical choice: Discriminating place and voice. *Perception and Psychophysics*, *43*, 21–30.

Whalen, D.H. (1983). The influence of subcategorical mismatches on lexical access. *Haskins Laboratories: Status Report on Speech Research*, *SR-73*, 1–15.

Wilson, R.H., & Carhart, R. (1971). Forward and backward masking: Interactions and addivity. *Journal of the Acoustical Society of America*, *49*, 1254–1263.

Appendix: list of test items for lexical decision tasks for Experiments 1 and 2

Related prime	*Unrelated prime*	*Target*
king	bell	queen
cat	fort	dog
cake	seed	pie
cape	meat	shawl
cane	sheet	stick
code	joke	spy
toast	size	bread
top	nut	bottom
test	rag	exam
tack	thief	pin
tall	choice	short
toe	hill	foot
tame	fan	wild
team	rain	coach
page	week	book
peace	hip	war
pine	seal	forest
pace	gift	step
pal	yawn	friend
poem	fog	verse
pill	race	drug
palm	job	hand
pear	jet	fruit
pit	sound	mine
pain	niece	hurt
pet	mesh	hug