Journal of Experimental Psychology:
Human Perception and Performance
1995, Vol. 21, No. 5, 1230–1235

Copyright 1995 by the American Psychological Association, Inc.
0096-1523/95/$3.00

Lexical Effects on Phonetic Categorization: The Role of Stimulus Naturalness and Stimulus Quality

Martha W. Burton
Pennsylvania State University

Sheila E. Blumstein
Brown University

A series of experiments was conducted to determine whether the effects of lexical status on phonetic categorization were influenced by stimulus naturalness (replicating M. W. Burton, S. R. Baum, & S. E. Blumstein, 1989, who manipulated the intrinsic properties of the stimuli) and by stimulus quality (presenting the stimuli in white noise). The experiments compared continua varying in voice onset time (VOT) only to continua covarying VOT and amplitude of the burst and aspiration noise in no-noise and noise conditions. Results overall showed that the emergence of a lexical effect was influenced by stimulus quality but not by stimulus naturalness. Contrary to previous findings, significant lexical effects failed to emerge in the slower reaction time ranges. These results suggest that stimulus quality contributes to lexical effects on phonetic categorization, whereas stimulus naturalness does not.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

A number of researchers have reported lexical effects on the phonetic categorization of speech (Connine & Clifton, 1987; Ganong, 1980; Miller & Dexter, 1988). These results have been cited in support of interactive models of speech processing such as the TRACE model in which higher level information influences lower levels of processing, in this case phonetic categorization (McClelland & Elman, 1986). Evidence for lexical effects on phonetic categorization have been typically demonstrated by shifts in the phonetic identification functions of listeners. For example, Ganong (1980) presented listeners with continua varying in voice onset time (VOT). The endpoints of these continua ranged from words to nonwords (e.g., *dash–tash, dask–task*). Listeners tended to label the ambiguous stimuli as words. Thus, the results showed more voiced responses in the boundary region for the continua in which the voiced endpoint was a word, and more voiceless responses when the voiceless endpoint was a word. Other studies have replicated this boundary shift (Connine & Clifton, 1987; Fox, 1984; Miller & Dexter, 1988). Nevertheless, in a recent study, McQueen (1989) found that with VOT continua, the obtained boundary shifts were small and highly variable. Within some blocks of trials, he found no evidence of a lexical shift. Furthermore, when a control nonword–nonword continuum was eliminated from the stimulus set, the lexical shift disappeared.

Pitt and Samuel (1993) have considered in detail the reliability of the lexical effect. Overall, they found that in 40 out of 55 data sets reported in the literature, there were significant lexical shifts; for the remaining 15 data sets, there were trends toward a lexical effect. As they point out, there are a number of procedural factors that affect the emergence of significant lexical effects. For example, a significant factor seems to be whether the stimuli are presented in a blocked or mixed design. The lexical effect seems much less likely to appear in designs blocked by continua than in mixed designs. Other factors identified as influencing the emergence of the lexical effect include the stability of the endpoint stimuli and the phonetic position of the manipulated segment. Those continua that have few ambiguous stimuli seem less likely to show lexical effects. In addition, lexical effects are less likely to emerge when the phonetic manipulation is in word initial position compared with word final position. Thus, an important consideration in evaluating the evidence for lexical effects is understanding the conditions under which the lexical effect emerges.

Burton, Baum, and Blumstein (1989) suggested an additional limitation on lexical effects in phonetic categorization. They showed that the emergence of the lexical effect depends on the intrinsic acoustic structure of the test stimuli. Using two continua that varied in VOT, one ranging from *duke* to *tuke* and the other from *doot* to *toot*, they showed a significant lexical effect. In addition, similar to previous results, the effect emerged in the slower reaction time ranges (Fox, 1984; Miller & Dexter, 1988). However, when the stimulus continua more closely approximated natural speech by covarying VOT and the amplitude of the burst and aspiration noise, the lexical effect disappeared. Moreover, when the data from the continua varying VOT and amplitude were divided into reaction time ranges, the lexical

---

Martha W. Burton, Department of Psychology, Pennsylvania State University; Sheila E. Blumstein, Department of Cognitive and Linguistic Sciences, Brown University.

This research was supported in part by National Institutes of Health Grant DC00142 to Brown University. Many thanks are due to Arthur Samuel for helpful comments and discussions on earlier versions of this article.

Correspondence concerning this article should be addressed to Martha W. Burton, 415 Moore Building, Department of Psychology, Pennsylvania State University, University Park, Pennsylvania 16802.

effect failed to appear in any of the reaction time ranges. Burton et al. (1989) argued that top-down processing effects may only emerge when the acoustic parameters of the stimuli are impoverished. However, in a recent reanalysis of the Burton et al. stimuli, it was found that the stimulus continuum covarying VOT and amplitude of the burst and aspiration noise did not vary parametrically as originally reported. That is, although the amplitude values were designed to increase parametrically from 49.8 to 56.8 dB, they instead varied from 46.8 to 50.3 dB, but not systematically across the continuum. Thus, because the stimuli did not comprise a continuum of values as originally reported, it is not clear whether the claims made by Burton et al. regarding the role of acoustic structure on the lexical effect can be replicated.

One goal of the current study is to replicate the original Burton et al. (1989) experiment. In addition, we intend to extend this research by investigating the role of stimulus degradation in the emergence of the lexical effect. There are two ways in which stimuli may be degraded or impoverished. First, the stimuli may vary in the extent to which they contain all of the parameters typically found in natural speech. That is, they may contain a reduced set of acoustic parameters or a conflicting set of acoustic parameters characterizing a particular phonetic dimension. We refer to these types of impoverished stimuli as a change in *stimulus naturalness.* The Burton et al. study, for example, showed that with increased stimulus naturalness, the lexical effect disappeared. Stimuli may also be impoverished when they are presented in noise. The normal listening situation is rarely optimal, and a decrease in the signal to noise (S/N) ratio will affect the perception of the stimulus set even though the acoustic parameters characterizing a particular phonetic dimension may still be present in the stimulus set. We refer to this type of manipulation as a change in *stimulus quality.*

A study by McQueen (1991) has explored the role of stimulus quality in the categorization of word final fricatives (e.g., *fish–fiss, kish–kiss*). He compared stimuli that were low-pass filtered to stimuli that were not. It was anticipated that word final position would be more likely to show lexical effects because by the end of the word, more phonetic and thus more lexical information has accumulated. However, the results showed that the lexical effect emerged in final position only when stimulus quality had been reduced and only in the fastest reaction time range.

Pitt and Samuel (1993) investigated the effect of degrading stimuli in both word initial and word final position. In one of their manipulations, they added noise to their stimuli by changing the sample values of digitized speech by a random amount. In both initial and final position, they found significant lexical shifts in the clear and in noise in the /g/–/k/ and /b/–/m/ continua. However, the magnitude of the lexical effect was only significantly larger in noise than in the clear condition in the /b/–/m/ continua in initial position. Interestingly, they found no significant lexical shift in the /d/–/t/ continua (ranging from *deep* to *teep* and *deach* to *teach*) with or without noise in initial position. They conclude that the addition of noise did not seem to have any systematic effect on the emergence of the lexical effect. Thus, results of previous studies exploring the role of stimulus quality have resulted in conflicting claims. McQueen (1991) argued that lexical effects depend on the acoustic quality of the stimulus, whereas Pitt and Samuel (1993) did not find any systematic effect of stimulus quality.

The role of stimulus quality in lexical effects has implications for models of speech perception. If stimulus quality plays a role in the emergence of lexical effects, then models of speech perception must provide mechanisms to account for these effects. For example, if lexical effects only occur when stimulus quality is degraded, then constraints on when higher level information can influence lower levels of processing must be incorporated into the model.

Two models that are distinguished by whether they allow interaction between the lexical and phonemic level are the TRACE model (Elman & McClelland, 1986; McClelland & Elman, 1986) and the autonomous race model (Cutler, Mehler, Norris, & Segui, 1987; Cutler & Norris, 1979). In the TRACE model, there are three levels: the feature level, the phoneme level, and the lexical level. There are bidirectional connections between each level that are facilitatory. In addition, there are inhibitory connections within each level (e.g., between phoneme nodes). In TRACE, the lexical effect occurs because of greater facilitation from the word level to the phoneme level for word responses than for nonword responses. The facilitation from the interaction between levels builds over time. Thus, the TRACE model predicts that in initial position larger lexical effects should occur in the slower reaction time ranges.

In contrast, the autonomous race model does not allow feedback from the lexical level to the phoneme level. Instead, this model claims that responses can be based on either prelexical information or the phonological codes associated with the lexical level. Processing on the prelexical and lexical level is parallel. The participant's response is based on whichever processing route finishes first. Task parameters and demands determine which level will win the race. Fast responses may be based on the prelexical route before the entire signal is heard, whereas for slow responses it is possible that the participants will have more lexical information available and the lexical route may win.

Although both models predict that lexical effects in initial position should emerge in the slower RT ranges, McQueen (1991) argued that the autonomous model is easier to modify to account for stimulus quality effects because the strategic mechanism necessary for variables such as stimulus quality are already built into the model. However, at present, this mechanism is only descriptive in nature. At present, TRACE does not have any mechanism for constraining the interaction between lexical and phonemic levels of processing.

In the present study, we sought to determine the extent to which stimulus naturalness and stimulus quality play a role in lexical effects. To that end, we created new stimuli to replicate the original Burton et al. (1989) experiment comparing stimuli that varied in VOT to those that covaried in VOT and amplitude. We used the same methods and procedures as in the earlier study. To explore the role of stimulus quality, we compared performance on stimuli pre-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

sented in noise and no-noise conditions. If degradation of stimuli quality does result in the emergence of the lexical effect, then even those stimuli that are unlikely to show a lexical effect without noise should show a lexical effect. On the other hand, if even with the addition of noise lexical effects do not emerge, it suggests a limited role for stimulus quality in the lexical effect.

Method

Stimuli

Four series of stimuli were constructed from different tokens of *duke, tuke, doot,* and *toot* recorded by the same speaker of American English as in the Burton et al. (1989) study. In the first series, the stimuli varied only in VOT; in the second series, VOT and amplitude of the burst covaried; and in the third and fourth series, the continua were embedded in noise. For the VOT continua, the tokens of *duke, tuke, doot,* and *toot* were computer edited following the methodology described in Burton et al. The stimuli were digitized on a PDP 11-34 computer at a sampling rate of 10.0 kHz with a 4.5 kHz low-pass filter and 10-bit quantization. They were then transferred to a VAX Station II and converted to 12-bit quantization for computer editing.

From measurements of the duration between zero crossings of the vowel in the waveform of the *duke* token, step sizes between tokens in the series were determined. The continuum steps ranged from 3.4 ms to 5.6 ms. The burst and original VOT of the [d] and a portion of the vocalic segment were replaced with a section of the burst and aspiration noise of the naturally produced [t] of the *tuke* stimulus such that as VOT increased, the vowel duration decreased. The continua consisted of 12 stimuli of equal duration with VOT ranging from 14.8 ms at the voiced end of the continuum to 64.6 ms at the voiceless end of the continuum. The amplitude of the burst and aspiration noise was measured for the *tuke* token using a 6-ms half-hamming window with a Patterson filter and preemphasis and was determined to be 54.4 dB (where $\text{dB} = 20 \log_{10}\text{RMS}$). This amplitude value was maintained across the continuum. Table 1 displays the VOT and vowel durations of the stimuli.

The consonant–vowel portions described above were used in both the *duke–tuke* and *doot–toot* continua. To create the final consonants for the *duke–tuke* and *doot–toot* continua, final consonants were excised starting at the onset of the closure interval and including the burst release from a token of *duke* and a token of *doot*. The final [k] or [t] was then appended to the consonant–vowel portions of each member of the continuum. Thus, two series of stimuli varying in VOT, one ranging from *duke–tuke* and the other from *doot–toot*, were created.

In the second series of continua, the stimuli were the same as those used in the VOT continua except that the amplitude of the burst and aspiration was systematically varied across the continuum. First, the amplitude of the burst and aspiration of the *duke* stimulus was measured and was determined to be 46.0 dB. The *tuke* token was measured to be 54.4 dB. The amplitude of the burst and aspiration was maintained for these two endpoint stimuli. For the rest of the stimuli on the continuum, the amplitude of the burst and aspiration was interpolated between the endpoint values such that there was a 0.76 dB difference in amplitude between adjacent stimuli. The appropriate scaling factor (sf), which would successively add 0.76 dB to each stimulus, was calculated using the following equation: $\text{sf} = e^{[(\ln 10)(\text{dB change})/20]}$. All other aspects of stimulus creation were identical to those of the VOT continua.

Two additional series of continua were created by adding noise to the VOT continua and to the continua varying in both VOT and amplitude. Pilot work suggested that a signal-to-noise ratio of +10.5 dB would degrade the perception of the endpoint stimuli but still allow participants to categorize the stimuli with a high degree of accuracy (better than 75%). To add the white noise to the test stimuli, we used the following procedure: The amplitude of the vowel in the test stimuli was determined by placing a rectangular window over 10 glottal pulses of the vowel and computing the root mean square energy converted to decibels. A white noise segment of equal duration to the stimulus was generated by a Grason-Stadler 901B noise generator such that it was 10.5 dB less than that of the measured amplitude of the vowel in the test stimuli. The white noise segment was then summed with each stimulus to create test continua with an S/N ratio of +10.5 dB.

It was also determined in pilot work that some participants were unable to identify the final consonants of the stimuli in noise. Because it was critical that participants perceive the final consonant (for determining whether the test stimulus was a word or not and thus for the lexical effect to emerge), before adding white noise to the stimuli the amplitude of the closure duration and release of the final stops [k] and [t] were scaled by a factor of 1.6 and appended to the original consonant–vowel portions of the stimuli. With these adjustments, the mean identification of the endpoint stimuli was 93% (with the lowest identification score of 1 participant at 80%), allowing participants to categorize the stimuli along the continua accurately including the final consonants. Thus, two continua varying in VOT and amplitude of the burst and aspiration were created with an S/N ratio of +10.5 dB.

Stimuli were then transferred to an IBM PC-AT for playout to participants. Each item on the continuum was presented 20 times in random order for a total of 240 stimuli per voicing series. There was a 3-s interval between stimuli and a 6-s interval between blocks of 10 items.

Table 1
Voice Onset Time (VOT) and Vowel Durations for Stimuli for All Continua

Stimulus	VOT	Step-size	Vowel duration
1	14.8		179.7
2	19.9	5.1	175.4
3	24.9	5.0	170.4
4	29.4	4.5	165.9
5	32.8	3.4	162.5
6	37.0	4.2	158.3
7	40.6	3.6	154.7
8	45.5	4.9	149.8
9	49.9	4.4	145.4
10	54.7	4.8	140.6
11	59.0	4.3	136.3
12	64.6	5.6	130.7

Participants

Fifty-six students at Brown University, all native speakers of American English, were paid for their participation in this experiment. None reported any history of hearing impairment, and all were naive to the purpose of the experiment. Twelve participants completed the four series of continua. For the noise VOT and amplitude continua, 20 participants were tested; however, 8 were eliminated on the basis of their performance on a posttest (see below).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Procedure

Participants heard one of the four series of continua, either the no-noise VOT continua, the no-noise VOT and amplitude continua, the noise VOT continua, or the noise VOT and amplitude continua. The presentation of the *duke–tuke* and *doot–toot* series of continua was blocked with order of presentation counterbalanced across participants. Thus, half of the participants listened to the *duke–tuke* series first, whereas the other half heard the *doot–toot* series first. Before the beginning of each test series, participants were presented with 20 practice items that included the 12 members of the particular continuum presented in order, followed by 8 endpoint and boundary stimuli in random order.

Participants were tested in groups of one to three and presented with stimuli over Sony MDR-2V headphones at a comfortable listening level. They were instructed to decide whether the first sound of the syllable that they heard was *d* or *t* and to press the appropriate response button, which was labeled *d* or *t*, as quickly as possible. The position of the response buttons was counterbalanced across participants and continua. Phonetic identification responses and reaction times were recorded by the IBM PC-AT that controlled the experiment.

For the noise continua, the procedure was the same as the no-noise continua with the following exception. Subsequent to presentation of the test series, participants were given a posttest that consisted of the endpoint stimuli from each continuum presented 10 times. Participants were asked to identify the stimuli as *duke, tuke, doot,* or *toot.* They indicated their responses by pressing one of four buttons that were labeled with orthographic representations of the syllables, *duke, tuke, doot,* and *toot.* This posttest was administered to ensure that participants could accurately perceive the endpoint stimuli.

Results

Before analysis of the identification functions, the percentage correct on endpoint stimuli from the posttest was scored for the participants in the noise conditions. For all of the participants in the VOT continuum, the mean percentage correct across endpoints was 80% or above. In the series of continua varying in VOT and amplitude, the mean score across endpoints was less than 60% for 7 of the 20 participants in that condition. Another participant failed to hear any of the *tuke* endpoint stimuli. These 8 participants were eliminated from data analysis. The remaining 12 participants in that condition correctly identified the endpoints at a rate of 80% or above.

The mean crossover boundary corresponding to 50% voiced responses was computed for each continuum for each participant by transforming the results to *z*-scores and performing a linear regression analysis. The data from all continua were analyzed together to evaluate the effects of stimulus quality and stimulus naturalness on the emergence of the lexical effect. Thus, a comparison of overall boundary values across all series of continua was made. Table 2 shows these results. A three-way analysis of variance (ANOVA) with two between-variables, noise (no-noise vs. noise) and acoustic parameter (VOT vs. VOT/AMP [amplitude]) and one within-variable, continuum (*duke–tuke* vs. *doot–toot*), was performed. Results showed a marginally significant effect for continuum, $F(1, 44) = 2.88$, $p < .10$, indicating a small lexical effect. There was a 0.15 stimulus boundary shift, which corresponded to 0.53 ms. In addition, there was a marginally significant interaction between noise and continuum, $F(1, 44) = 3.13$, $p = .08$, due to a larger lexical effect in the noise conditions (0.30 stimulus shift) than in the no-noise conditions (−0.01 stimulus shift). No other main effects or interactions were significant.

Table 2
Phonetic Boundary for Duke-Tuke *and* Doot-Toot *and Magnitude of Boundary Shifts Across the Series of Continua*

Condition	Duke–Tuke	Doot–Toot	Boundary shift
Without noise			
VOT	6.60	6.53	0.07
VOT/AMP	6.45	6.53	−0.08
With noise			
VOT	6.87	6.67	0.20
VOT/AMP	7.02	6.62	0.40

Note. VOT = voice onset time; AMP = amplitude.

Following previous studies (Burton et al., 1989; Fox, 1984; Miller & Dexter, 1988), the responses were divided into reaction time ranges to determine whether the lexical effect emerged in particular reaction time ranges. To that end, we divided the reaction time data into thirds for each participant. Boundary values were calculated for each of the three reaction time ranges. If no responses were recorded for a particular stimulus, no value was entered. As in Miller and Dexter (1988), participants with nonmonotonic functions in the boundary area were eliminated from the reaction time range analysis. In each of the two types of no-noise continua, 2 participants were eliminated. In the VOT noise continua, 5 participants were eliminated and three were eliminated in the continua varying in VOT and amplitude in noise. Results are presented in Table 3.

A four-way ANOVA was performed with two between-subjects variables, noise (no-noise vs. noise) and acoustic parameter (VOT vs. VOT/AMP), and two within-subjects variables, range (fast vs. medium vs. slow) and continuum (*duke–tuke* vs. *doot–toot*). Most important, there was a main effect of continuum, $F(1, 32) = 7.59$, $p < .01$. For the *duke–tuke* continuum, the mean boundary value was 6.62, whereas for *doot–toot,* the mean boundary value was 6.38, a difference of 0.86 ms. Thus, there was a significant shift in the boundary depending on lexical status of the endpoints. None of the interactions that would indicate effects of stimulus naturalness and stimulus quality on the lexical effect were significant. There was a nonsignificant Acoustic Parameter × Continuum interaction ($F < 1$). A significant interaction would have indicated an effect of stimulus naturalness. The mean difference in boundary shift was 0.24 for the VOT continua, whereas for continua varying in VOT and amplitude, it was 0.27. Thus, there was little difference in the boundary shift between the VOT continua and the continua varying in VOT and amplitude. Moreover, there did not appear to be a significantly larger lexical effect in noise than in no-noise, which was indicated by the lack of a significant Continuum × Noise interaction, $F(1, 32) =$

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Phonetic Boundaries at Three Reaction Time Ranges and Boundary Shifts in Each of the Continua

Condition	Fast			Medium			Slow		
	Duke–Tuke	Doot–Toot	Boundary shift	Duke–Tuke	Doot–Toot	Boundary shift	Duke–Tuke	Doot–Toot	Boundary shift
Without noise									
VOT	6.07	5.90	0.17	6.45	6.40	0.05	6.72	6.45	0.27
VOT/AMP	6.36	5.99	0.37	6.40	6.24	0.16	6.39	6.55	−0.16
With noise									
VOT	6.68	6.51	0.17	6.82	6.51	0.31	6.86	6.40	0.46
VOT/AMP	6.99	6.66	0.33	6.97	6.52	0.45	7.06	6.62	0.44

Note. VOT = voice onset time; AMP = amplitude.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

1.41, $p = .24$. Nonetheless, consistent with the overall analysis, there was a trend in the direction of an effect of stimulus quality. The stimulus shift difference in boundary values for the noise condition was 0.37, whereas it was 0.15 for the no-noise condition.

In addition, the four-way ANOVA revealed a number of other significant effects. However, none of these effects were related to the parameters of interest, that is, the effects of stimulus naturalness or stimulus quality on the lexical effect. Results showed a main effect of noise, $F(1, 32) = 4.29$, $p < .05$. The mean boundary value for no-noise was 6.32 compared with 6.71 for noise. Thus, more *t* responses occurred at the boundary in the no-noise continuum than in the noise continuum. There was also a main effect of range, $F(2, 64) = 4.44$, $p < .05$. The mean boundary value for the fast range was 6.36, for the medium range it was 6.51, and for the slow range it was 6.62. Newman-Keuls post hoc tests showed that the fast and slow ranges were significantly different. The main effect of acoustic parameter was not significant ($F < 1$), indicating that the boundary values did not differ depending on whether the participants were presented with the VOT or the continua varying in VOT and amplitude. Finally, there was a Range × Noise interaction, $F(2, 64) = 3.65$, $p < .05$, due to a shift in phonetic boundary as a function of reaction time range but only for the no-noise condition.

Discussion

This study explored the effects of stimulus naturalness and stimulus quality on phonetic categorization. Although the results showed an overall lexical effect, these effects emerged independent of stimulus naturalness. If stimulus naturalness were a critical parameter in the emergence of the lexical effect, there should have been a significant Continuum × Acoustic Parameter interaction. Thus, in contrast to the conclusions of Burton et al. (1989), stimulus naturalness does not seem to play a role in the lexical effect.

To further investigate the lack of replication of our previous results, we ran an additional experiment using the original VOT continua from the Burton et al. (1989) study. The stimuli and procedures were the same as Experiment 1 in Burton et al. except that 20 repetitions rather than 10 repetitions of the stimuli were presented to a new group of listeners. Results showed no lexical shift (−0.03) and were confirmed by a one-way repeated measures ANOVA. In addition, the results showed no significant differences in boundary values for any of the reaction time ranges. Thus, the effect of stimulus naturalness on the lexical effect in Burton et al. could not be replicated with new stimuli or the original stimuli.

When the stimuli from the no-noise condition were embedded in white noise, there appeared to be an increase in the magnitude of the lexical effect, although it appeared only as a trend in the overall analysis of the boundary values and failed to emerge in reaction time range analysis. The marginally significant interaction between noise and continuum in the overall analysis suggests that stimulus quality contributes in a small way to the emergence of lexical effects.

It is noteworthy that the lexical effect, although small (with an overall boundary shift of 0.53 ms), still emerged despite the fact that the stimulus conditions and parameters were constructed to minimize this effect. That is, stimuli were [d]–[t] continua presented in a blocked design with perceptually stable endpoints that were in initial position. Each of these factors have been reported to minimize or eliminate lexical effects (Pitt & Samuel, 1993). These results suggest that the lexical effect can be obtained even when conditions are weighted against it.

Contrary to the hypothesis that lexical effects would emerge (or if already present, be magnified) at the slower reaction time ranges (Burton et al., 1989; Fox, 1984; Miller & Dexter, 1988), no such effects were obtained in the current study. Nonetheless, as Table 3 shows, there seems to be a tendency in the means toward larger lexical effects in the slower reaction time range, particularly in noise. At present, we have no explanation for the failure to obtain significant effects in the slower reaction time ranges.

In addition to the lexical effect, the addition of noise to the stimuli resulted in an increase in the number of *d* responses at the boundary. This increase in *d* responses in noise suggests that the masking of the signal may affect the perception of [d] and [t] unequally in voicing continua. Voiceless consonants may be more affected by degrading stimulus quality with white noise than voiced consonants. The acoustic parameters that contribute to voiceless stop consonants, in particular, the burst and aspiration noise, are signaled in the high-frequency range, whereas the acoustic parameters that contribute to voiced stop consonants have greater low-frequency information. Because of this potential asymmetric effect of noise on [d] and [t], it may be more difficult for lexical effects to emerge in noise in voicing continua than in continua that vary other phonetic contrasts. Some evidence to support this possibility is that the only one of Pitt and Samuel's (1993) continua that showed even marginally larger lexical effects in noise than in no-noise was the [b]–[m] continua in which both endpoints are voiced.

The results of this study do not favor either the autonomous race model or the TRACE model (cf. Pitt & Samuel, 1993). However, the results present a challenge for the models to explain. Our results suggest that the models need to make provisions for larger lexical effects when stimulus quality is degraded. Whatever the model, there should be motivated mechanisms to predict when the lexical effect will and will not emerge.

In any case, listening to speech in noisy conditions used in this experiment may be analogous to what the listener faces in perceiving speech everyday. Listening to speech in a laboratory setting over headphones may not be a "normal" listening condition. Under such ideal listening conditions, it may be possible to rely on low-level acoustic information (i.e., prelexical information). In conditions of noise, however, participants may need to resort to lexical strategies in processing the stimuli, hence a larger lexical effect.

References

Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 567–575.

Connine, C. M., & Clifton, C. (1987). Interactive use of lexical information. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 291–299.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology, 19,* 141–177.

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 113–134). Hillsdale, NJ: Erlbaum.

Elman, J., & McClelland, J. (1986). Exploiting lawful variability in the speech waveform. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 360–380). Hillsdale, NJ: Erlbaum.

Fox, R. A. (1984). Effects of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 526–540.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 110–115.

McClelland, J. L., & Elman, J. L. (1986). Interactive processes in speech perception: The TRACE model. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Exploration in the microstructure of cognition* (vol. 2, pp. 58–121). Cambridge, MA: MIT Press.

McQueen, J. M. (1989). The use of lexical knowledge in phonetic categorization. In J. P. Tubach & J. J. Mariani (Eds.), *Eurospeech '89: Proceedings of the European Conference on Speech Communication and Technology* (vol. 2, pp. 581–584). Edinburgh, Scotland: CEP Consultants.

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

Miller, J. L., & Dexter, E. R. (1988). Effect of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 369–378.

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

Received December 28, 1990
Revision received July 18, 1994
Accepted September 30, 1994 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.