

Subcategorical phonetic mismatches and lexical access

D. H. WHALEN

Haskins Laboratories, New Haven, Connecticut

The place of phonetic analysis in the perception of words is unclear. While some theories assume fully specified phonemic strings as input, other theories assume that little analysis occurs. An earlier experiment by Streeter and Nigro (1979) produced evidence, based on auditorily presented words with misleading acoustic cues, that lexical decisions were based on mostly unanalyzed patterns, since word judgments were delayed by misleading information whereas nonword judgments were not. The present studies expand that work to a different set of cues, and to cases in which the overriding cue came first. An additional task, auditory naming, was used to examine the effects when the decision stage is less demanding. For the lexical decision task, misleading information slowed the responses, for both words and nonwords. In the auditory naming task, only the slower responses were affected. These results suggest that phonetic conflicts are resolved prior to lexical access.

The level of attention paid to the phonetic detail of auditorily perceived words is a matter of some controversy. The large literature on phonetically untrained listeners' extreme sensitivity to a multitude of cues (see, e.g., Darwin, 1976; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Pisoni, 1978, 1985; Repp, 1982) derives its force from the assumption that listeners actually use this sensitivity in normal speech perception. Yet many investigators have proposed an insensitivity when the mental lexicon is accessed (e.g., Klatt, 1979; Marslen-Wilson, 1984; Morton, 1969; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985).¹ These models tend to focus on the probabilistic nature of lexical access and thus allow for much greater latitude when dealing with the speech signal itself. In a sense, the labels for the two fields define what the subjects produce: speech perception or lexical access. The two experiments reported here provide some evidence for joining the two.

Although there have been many studies of the effects of phonological similarity on auditory lexical access, it is difficult to find examples of the reverse—namely, lexical effects on phonetic judgments. Ambiguous items tend to be heard as words rather than nonwords (Ganong, 1980), though this may be due to a nonperceptual strategy. In a test of speeded identification with similar stimuli, Fox (1984) found that only slow responses showed a lexical-

ity effect. He interpreted this to mean that it is possible for the phonetic decisions to be made without lexical input, but that if the response was not made soon enough, the lexicon would influence ambiguous decisions. Miller and Dexter (1988) found a similar lack of lexical influence when subjects were required to respond quickly, though the responses also showed that the subjects were responding as if the items presented were in fact very short. When the subjects were not able to respond quickly, lexical effects still emerged. Burton, Baum, and Blumstein (1989) also found a lexical effect in slow responses to ambiguous items, but only when the items consisted of synthesized syllables. When Burton et al. modified natural speech to produce a less schematic stimulus, the lexical effect disappeared altogether. McQueen (1991) also found lexical effects on an /s/-/f/ continuum only with degraded stimuli. The influence of the lexicon is made more conditional by the finding that the lexical entry needed for a nonword to act like a word can be established during the course of eight presentations in a story heard minutes before the perception takes place (Remez, Rubin, Katz, & Dodelson, 1985). None of these studies, of course, assessed speech processing in normal conversation, in which stimulus degradation may be relatively common. Still, these results in general suggest that phonetic decisions are autonomous with respect to the lexicon.

Perhaps the best example of the lexicon's apparently avoiding phonetic analysis is the study of mismatched transitions for stops in Streeter and Nigro (1979). When the transitions going into and out of a stop closure indicate different places of articulation, the perceived place almost always is that of the transitions after release. Using a variety of measures—but most relevantly, in a lexical decision task—they found that such mismatches affected judgments of words but not of nonwords. They interpreted this to mean that the lexicon is searched with a faithful,

The support of NIH Grant HD-01994 to Haskins Laboratories has made this research possible. Experiment 1 formed part of a dissertation (Whalen, 1982). Experiment 2 was presented at the 109th meeting of the Acoustical Society of America, Austin, TX, April 1985. André M. Cooper provided invaluable assistance in running Experiment 2. Helpful comments were provided by Alvin M. Liberman, Louis M. Goldstein, Arthur G. Samuel, Carol A. Fowler, Len Katz, and two anonymous reviewers. Correspondence should be addressed to D. H. Whalen, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695.

if sometimes degraded, representation of the speech input. The lack of an effect for nonwords would be due to a failure of an exhaustive search of (part of) the lexicon, after which a "nonword" judgment is made.

Similar mismatches, though, have been shown to be processed in more explicitly phonetic tasks (Martin & Bunnell, 1982; Whalen, 1982, 1984; Whalen & Samuel, 1985). In my work, the cues to the fricative place of articulation in the fricative noise and in the vocalic formant transitions were put into competition. It was found that the noise information for the alveolar and palatal fricatives (/s/ and /ʃ/) consistently overrode the formant information. Mismatches were not consciously detectable, but they affected the speed with which the subjects could make phonetic identifications. Here, it did not make any difference whether the overriding cue came before or after the other cue.

The present experiments explore the effect of these subcategorical phonetic mismatches on auditory lexical access. The first experiment is a lexical decision task, like that used by Streeter and Nigro (1979). The second uses the stimuli of the first experiment in an auditory naming task (cf. Tyler, Marslen-Wilson, Rentoul, & Hanney, 1988), in which the demands of the decision stage of lexical decision have been shown to be reduced (Balota & Chumbley, 1984, 1985). If the longer decision times for nonwords are due to a rechecking of possible lexical entries rather than the way the initial search is conducted, the decision times might obscure an effect of phonetic processing for the nonwords. The naming task could be expected to avoid such decision-related factors and, at least for the times that begin late enough for the phonetic mismatches to have been processed, to allow the effects of mismatches to appear.

In addition to the differences in lexical status, the items also varied in the location of the fricative, and thus in the order of the overriding and subordinate cues. If the temporal order of the cues determines their usefulness in lexical access, perhaps only subordinate cues that precede the overriding cues will affect the processing times (as assumed by Streeter & Nigro, 1979). However, the order of the cues does not affect phonetic decisions (Whalen, 1984), and so the same may be true of lexical access. The items also differed in whether or not the change between the two fricatives made a change in the lexical status. This manipulation results in some cases in which a "nonword" decision could be made without resolving the phonetic mismatch. For some theories, such as phonetic refinement theory (Pisoni et al., 1985), only as much phonetic detail is extracted as is needed for lexical identification. In that light, mismatches that would only distinguish between two nonwords or between two words would be expected not to be used during a lexical decision task, but might affect naming times. Thus the convergence of the two experiments should begin to illuminate the role of phonetic perception in lexical access.

EXPERIMENT 1

The first experiment examined the effect of subcategorical mismatches on the processes involved in the lexical decision task.

Method

Stimuli. In previous work on subcategorical mismatches, the fricatives /s/ and /ʃ/ were used exclusively (Whalen, 1984). So in the present work, items were selected to contrast those two fricatives, as in *mess* and *mesh*. Only monosyllabic items were used. Since few items with high printed frequency contain these two fricatives, the contribution of word frequency was minimized by selecting the words that had less than 50 occurrences in the Kučera and Francis (1967) corpus. For items that had homonyms or inflections, the total of the occurrences was less than 50.

Four parameters guided the choice of items: fricative (/s/ or /ʃ/), lexicality (word or nonword), location (initial fricative or final fricative), and *changeability*. This last category describes the lexical status of the item when the fricative identity is changed. The status either remains the same (as when *mess* goes to *mesh* or *gish* goes to *giss*) or changes (as when *soak* goes to *shoak* or *shilk* goes to *silk*). Within the 16 categories generated by these parameters, 6 items were selected, resulting in 96 stimuli (see the Appendix).

For several nonwords with final /s/, the pronunciation given could easily be a fast speech version of a final /st/ cluster. Thus *kass*, *goass*, and *duss* could be heard as shortened versions of *cast*, *ghost*, and *dust*. Several words (such as *bus* and *lass*) could also be heard as different words through the same process. This necessitated explicit instructions to the subjects that these were careful pronunciations, and that therefore these /t/s had to be pronounced in order to affect the judgment.

Of the four parameters, three have theoretical implications. Since mismatches affect both /s/ and /ʃ/, the fricative parameter generates no predictions and will thus be excluded from the analysis. Lexicality would be predicted to affect the processing of the mismatches, given that the most similar previous study (Streeter & Nigro, 1979) resulted in a difference between words and nonwords. The location of the fricative makes a difference in some theoretical perspectives (e.g., cohort theory; see Marslen-Wilson, 1984; Marslen-Wilson & Welsh, 1978), but not in others (e.g., phonetic refinement theory; see Pisoni et al., 1985). Changeability would be predicted by most theories to affect the processing, since most would assume that (at least for nonwords) information that would not change the lexical decision would not be fully processed.

For every test item, there were two filler items that did not contain /s/ or /ʃ/ outside of a cluster. These items were selected so that approximately the same distribution of phonemes occurred as in the test items, with the exception of /s/ and /ʃ/. For filler words, frequency of printed occurrence was matched to that of the test words by selecting two monosyllabic words with the same frequencies. Nonwords were formed from these through the change of a single phoneme. Some of these items were chosen because they contained the /st/ final cluster that made some of the test items problematic. This ensured that the subjects would hear some items with a released /t/, giving them experience with the careful pronunciation that would have been necessary to ensure that items like *kass* were treated as nonwords.

Twenty words and 20 nonwords were also selected along similar criteria, with the exception that some /s/ and /ʃ/ words were included (e.g., *sang*). These were used to familiarize the subjects with the task. Eight more filler items occurred as warm-up items at the beginning of the test sequence; they were eliminated from the analysis.

Procedure. A male speaker of American English recorded three randomized lists of all test and filler items. These recordings were input into a VAX-11/780 computer via the Haskins PCM system (Whalen, Wiley, Rubin, & Cooper, 1990) at a 20-kHz sampling rate. From these, a single token was selected for the test. Filler items were chosen on the basis of clarity of pronunciation, as judged by the experimenter. Test items were chosen in part on the same basis, but with the further constraint that the item that had the other fricative in it had to have the same duration in its components. That is, both the fricative noise and the remainder of the syllable had to have the same duration. Of course, measurements of gradual onsets and offsets, as are usually found with fricative noises, have a certain amount of inherent arbitrariness, but the amplitude contours appeared well matched too. In this way, the fricative noises could be transposed without changing the total duration of the items. The experimenter also verified that the mismatched items did not result in any overt auditory discontinuities.

For each test item, then, a version containing a mismatch was created by substituting the vocalic segment and remaining consonant(s) from the item that differed from it only in fricative identity. For example, the item *mess* would appear in one case in its original form, and in the other with the /m/ resonance and vocalic segment of *mesh*. Since the fricative noise determines the identity of the fricative in this case, both are perceived as *mess*. Each subject was to receive each version, so two randomizations were made, one with one version of an item, the other with the other. The order in which these two randomizations were presented was counter-balanced across subjects.

Stimuli were recorded onto one channel of an audio tape for binaural presentation to the subjects. Simultaneous with the onset of each item was a brief tone (inaudible to the subjects) that was recorded on the other channel of the tape. Three seconds of silence separated one stimulus from the next. The stimuli were presented over headphones (TDH-49) to the subjects, while the tone was used to start the clock for the reaction time. The subjects' judgments and times were recorded with the pressing of either a button labeled "word" or one labeled "nonword." If no response was made within 2.5 sec, the response was coded as "missing." The subjects were run either alone or with 1 or 2 other subjects.

Each session began with a description of the lexical decision task, with emphasis being put on the fact that these were careful pronunciations. The subjects were instructed to be as quick as possible, but to slow down if they began making too many mistakes. A small red light was lit for each item correctly categorized. After the practice, the subjects were asked if any further clarification was needed. Then came the first test sequence, which contained the 8 warm-up items followed by the 96 test items, half with mismatched transitions and half in their original form, randomized with the 192 filler items. Feedback was given during this portion as well. When this block was done, there was a brief break followed by the other tape, which contained the complementary version of each stimulus.

Subjects. The subjects were 22 males and 14 females who were native speakers of American English of varying varieties. They ranged in age from 18 to 45, and they had no reported hearing problems, except for 1 subject who had normal hearing in one ear and some impairment in the other. The subjects were also classified as being phonetically trained or not, divided equally between the males and the females. The phonetically untrained subjects were Yale undergraduates who were paid for their participation. The rest had received at least some training in phonetics and were colleagues from Haskins Laboratories or the Yale Linguistics Department, who donated their time. If the effect of mismatching were to show up only for those with training, we would have more evidence that focusing on the phonetic structure is necessary for these fine-grained analyses to be important. If all the subjects should turn out to respond similarly, then we could presume that the mismatches are detected without special attention.

Results

Both the error and the reaction time values were subjected to an analysis of variance with the factors expertise (phonetically trained or not), condition (first presentation or second), lexicality (word or nonword), position (fricative initial or fricative final), changeability (whether the change in fricative results in a change of lexical status or not), and appropriateness (whether the transitions were original or mismatched). Expertise was a between-subjects factor for the subject analysis, and lexicality, changeability, and position were between-items factors for the item analysis. In the remainder of this section, the *F* ratios for the subject analysis always precede those for the item analysis.

Although the primary measure of interest is reaction time, the pattern of errors is of some value as well. Overall, the error rate on test items was 8.0%, which was in line with the rate of 7.8% for the filler items. The rate for test items was higher for the first tape (9.3%) than for the second (6.7%) [$F(1,34) = 12.67, p < .01$; $F(1,88) = 17.81, p < .001$]. Because of the disparities between the changeable and the nonchangeable words, none of the error differences were significant in both the subject and the items analyses. Words were misclassified more frequently than nonwords [9.6% vs. 6.4%; $F(1,34) = 12.64, p < .01$; $F(1,88) = 2.19, n.s.$]. This pattern will be examined later in conjunction with the reaction time results for the lexicality factor. Whatever trend might have occurred in the mismatching condition favored the mismatched versions, which were misclassified somewhat less frequently than the matched ones (7.8% vs. 8.2%; both *F*s < 1).

For the reaction time analysis, times for incorrect responses were replaced with estimates based on a regression using all the correct responses. Since the variances tended to increase as the time increased, the times were transformed into speeds before the regression and subsequent analysis of variance were performed. In reporting the values, however, speeds are retransformed into times to make comparisons with other research more convenient (see Table 1).

The primary factor of interest, appropriateness of transitions, showed significantly faster times for items with appropriate transitions than those with inappropriate ones [884 vs. 895 msec; $F(1,34) = 10.02, p < .01$; $F(1,88) = 5.57, p < .05$]. Although no interactions between this factor and any others reached significance in both the subject and the item analyses, three were significant in the subject analysis. For these items, then, the delay was larger for words than for nonwords [16 vs. 5 msec; $F(1,34) = 6.18, p < .05$; $F(1,88) = 1.58, n.s.$]. (See Figure 1.) The delay was larger for items with final fricatives than for those with initial fricatives [18 vs. 3 msec; $F(1,34) = 6.25, p < .05$; $F(1,88) = 2.99, n.s.$]. Finally, there was an interaction of appropriateness, lexicality, position, and changeability [$F(1,34) = 5.89, p < .05$; $F(1,88) = 2.38, n.s.$]. This interaction does not have any obvious interpretation.

Table 1
Means for Lexical Decision Times (Experiment 1)

Transitions	Words				Nonwords			
	Initial		Final		Initial		Final	
	Changeable	Stable	Changeable	Stable	Changeable	Stable	Changeable	Stable
Condition 1								
Appropriate	906	835	917	894	951	920	924	833
Inappropriate	912	847	934	916	951	920	911	873
Condition 2								
Appropriate	882	810	886	863	928	907	906	824
Inappropriate	881	838	906	885	935	910	894	832

Note—Times given in milliseconds.

Responses were faster in the second condition [901 msec in the first, 879 msec in the second; $F(1,34) = 11.62, p < .01$; $F(1,88) = 64.28, p < .01$]. Because of the nature of the experiment, it is not possible to tell whether this is a practice effect, a repetition priming effect, or a combination of the two. In any event, condition did not enter into any interactions that were significant in both analyses.

The trained subjects were faster than the untrained [842 vs. 943 msec; $F(1,34) = 8.94, p < .01$; $F(1,88) = 724.49, p < .01$]. It is unlikely that phonetic training per se increases one's speed, so this is probably due to a difference in motivation or in familiarity with reaction time paradigms. Expertise entered into only one interaction, with location [$F(1,34) = 14.24, p < .01$; $F(1,88) = 13.07, p < .01$]. Although both trained and untrained subjects were faster in responding to fricative final items, the effect was more pronounced for the trained subjects (a difference of 53 msec for the trained vs. 36 msec for the untrained). This may reflect a greater sensitivity to cues in the offset of the fricative noises indicating a lack of a following stop. If the subject knew that there was no stop to come, it would make for a faster decision.

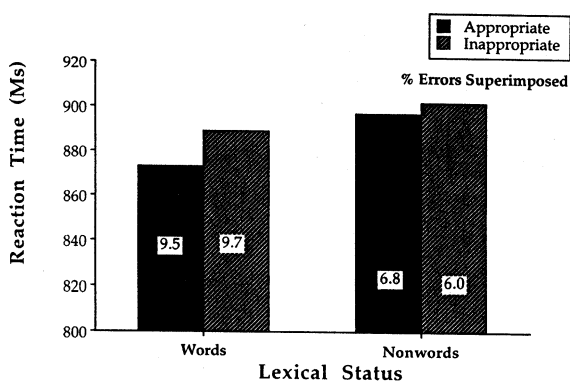


Figure 1. Effects of appropriateness of transition on lexical decisions for words and nonwords. Values represent averaged speeds retransformed into times.

Although nonwords were categorized more slowly than words (900 vs. 881 msec), only the subject analysis showed it to be a significant difference [$F(1,34) = 9.35, p < .01$; $F(1,88) = 1.81, n.s.$]. Although the size of the difference is smaller than that of other results in the literature, the composition of the words in the present experiment helps explain the discrepancy as well as the lack of significance for the item analysis. Overall, a slight difference in stimulus duration favored the nonwords (575 msec average duration for words, 564 for the nonwords). (Duration was significantly correlated with average reaction time, $r = .34, p < .05$.) This difference is not large enough to explain the lack of the usual effect, however. Instead, properties of the words seem to be responsible. Half of the words in this study were deliberately chosen to differ only in having /s/ or /f/ in one location. This increased the difficulty of some decisions, as can be seen in the longer time taken for items that changed lexical status with the change in fricative (894 vs. 886 msec for the nonchanging items), though changeability was only significant for the subject analysis [$F(1,34) = 9.65, p < .01$; $F(1,88) < 1$]. The interaction of changeability and lexicality, though, was significant in both analyses [$F(1,34) = 177.52, p < .01$; $F(1,88) = 13.17, p < .01$]. Although changeability did not affect the word judgments (863 vs. 860 msec), there was a large effect for the nonwords (927 msec for changeable nonwords, 876 msec for the stable ones; there was no speed/accuracy trade-off, since the changeable nonwords yielded a 7.8% error rate compared with 5.0% for the stable ones). This may be a combination of the difficulty of the changeable nonword decisions and the potentially smaller neighborhoods of the stable nonwords (e.g., *froose/froosh*). Similarly, the speed obtained for the changeable words may have been at the expense of accuracy, since words were more frequently misclassified. (The contribution of the item *deuce* may be excessive here, since it alone accounted for 9.2% of all errors.)

Items with initial fricatives elicited slower times than those with final fricatives [914 vs. 868 msec; $F(1,34) = 165.54, p < .01$; $F(1,88) = 13.37, p < .01$]. This may be due to differences in familiarity in the items, and, in

any event, it did not have specific theoretical impact as a main effect. The interactions, which might have been predicted by cohort theory, were not significant.

Discussion

Subcategorical phonetic mismatches were found to affect the decision times for distinguishing words and nonwords. Although the overall analysis did not reveal a difference between words and nonwords in this regard, there are several indications that the effect may not hold for nonwords. There are similar weak indications that there may be a difference in the treatment of these mismatches, depending on whether they occur early in the item or late. For the most part, though, these mismatches seem simply to impair the lexical decision process wherever they occur.

Such a ubiquitous effect suggests that phonetic analysis resolves the conflict within the stimulus before searching the lexicon. Just the opposite interpretation was made by Streeter and Nigro (1979), who found no effect with nonwords. Thus they concluded that "stimulus degradation" in the mismatched cases is not resolved prior to lexical search, and that the nonwords showed no effect because *nonword* judgments are assumed to be the failure of an exhaustive search of the lexicon. The appearance of a discrepancy between these two sets of results may be misleading. If Streeter and Nigro had analyzed their results in one analysis rather than in separate ones for words and nonwords, they might have obtained the results reported here. Indeed, if the present results are analyzed separately, there is a significant effect for words but not for nonwords. The lack of an interaction indicates that nonword judgments are affected by phonetic mismatches. At the least, the present results do not argue strongly for a separate explanation of the nonword results. Indeed, even within Streeter and Nigro's account, it would seem that they would predict an ever increasing effect of mismatching with increasing size of the lexical search—that is, with either nonwords or infrequent words. Although they describe only the final match as being affected by the degradation of the input, surely every comparison would be affected. So if the search was exhaustive (of all or part of the lexicon), it would have even more of these slower comparisons than would a successful match. If instead we assume that the mismatch is corrected initially, we might expect that such a delay would be less and less of a factor as the reaction time increased, which the nonword judgments in fact show.

The fact that the stable nonwords with final fricatives still showed an effect of the transitions is difficult to accommodate in the cohort theory. Only one of the 12 items had any potential word candidates remaining after the final /s/ or /ʃ/: *duss* still has the potential for resulting in several words containing *dust* or *dusk*. For the rest (assuming that none of the subjects had *couchee* as a possible continuation of *koosh*), either fricative would terminate the lexical search. Nonetheless, the appropriateness of the transitions had a substantial (24-msec) effect on the judgments for these items. It appears that the fricative speci-

fied was determined by the subjects, even though either fricative would have triggered a nonword response in the cohort theory.

Further elaboration of these results will be postponed until the general discussion, when the results from the auditory naming task will also be available.

EXPERIMENT 2

In the second experiment, the same stimuli were used in a different task—an auditory naming task. Subjects heard the words and nonwords and were to repeat them as quickly as possible. The naming task should be less subject to the complications of the decision stage of the lexical decision task while still containing a component of lexical access (Balota & Chumbley, 1984, 1985).

Method

Stimuli. The stimuli were the same as in the first experiment, but two additional permutations of the sequences were used. It appeared, from informal observation, that the two randomizations used in the first experiment resulted in some list effects. To minimize the impact of such effects on the parameter of most interest—namely, the appropriateness of the transitions—two more versions of each randomization were made. The order of the items was identical in these new versions, but the appropriateness of the transition was changed for each item. In this way, whatever list effects were present would now be distributed evenly across the two levels of appropriateness.

Procedure. The subjects heard the stimuli in the same situation as they did in the first experiment. Their responses, though, were made by repeating the item into a microphone placed in front of them. When the acoustic energy reached a critical level (which was set individually for each subject), an impulse was sent to the reaction timer. An Atari 800 computer collected the times in this experiment. The speech was also transmitted to another set of earphones so that the experimenter could listen and note incorrect responses. Although the exact incorrect response was written, no theoretical questions depended on them, so they will not be reported.

The acoustic envelopes of different classes of phones are quite different, an effect that can bias the measured reaction times when items in one experimental category have different onset characteristics from those in another category. This was certainly the case here, where the items with initial fricatives always had gradual onsets, whereas the items with final fricatives tended to have initial stops with their abrupt onsets. Although the most important factor, the appropriateness of the transitions, was located within a single phonetic sequence for each item, the other comparisons needed some adjustment for the onset differences. This was accomplished by running the stimulus tape and letting it be its own subject. The stimulus itself was fed into the line that generated the response triggers. The result was a number giving an average onset delay, which was then subtracted from each of the subject responses.

Subjects. A different set of 36 subjects participated in the second experiment. Since phonetic training did not affect anything other than the overall speed of the subjects, an unbalanced selection was made. Eleven colleagues from Haskins Laboratories (6 female, 5 male) volunteered their time. The other 25 (14 female, 11 male) were Yale undergraduates who were paid to participate. None reported any hearing difficulty.

Results

The overall error rate was 6.2%, slightly less than the 8.3% for the filler items. Appropriateness of transition

Table 2
Means for Auditory Naming Times (Experiment 2)

Transitions	Words				Nonwords			
	Initial		Final		Initial		Final	
	Changeable	Stable	Changeable	Stable	Changeable	Stable	Changeable	Stable
Condition 1								
Appropriate	382	361	464	456	367	393	444	427
Inappropriate	394	366	435	458	393	404	458	449
Condition 2								
Appropriate	302	296	405	403	290	332	393	392
Inappropriate	301	301	397	410	308	333	402	395

Note—Times given in milliseconds.

had no effect at all on this value (6.2% for each). Nonwords were mispronounced significantly more often than words [8.6% vs. 3.8%; $F(1,35) = 54.38, p < .01$; $F(1,88) = 13.10, p < .01$]. (Although there were cases of nonwords being pronounced as words, there were just as many cases of words and nonwords being pronounced as other nonwords.) Items with initial fricatives were mispronounced significantly more often than those with final fricatives [8.0% vs. 4.4%; $F(1,35) = 26.36, p < .01$; $F(1,88) = 7.76, p < .01$]. There was a nonsignificant improvement in the second condition (6.5% for the first, 5.9% for the second).

For the analysis of the naming times, it was found that the variance was fairly homogenous without a transformation, so none was used. The times presented do, though, have the "microphone response time" subtracted from them. The means are presented in Table 2. In the remainder of this section, the subject analysis precedes the item analysis.

Appropriateness of transition had an effect only in the subject analysis in this experiment [382 vs. 388 msec; $F(1,35) = 5.09, p < .05$; $F(1,88) = 1.50, n.s.$]. There was also an interaction with Lexicality in the subject anal-

ysis [$F(1,35) = 5.91, p < .05$; $F(1,88) = 2.04, n.s.$]. But in this case, as can be seen in Figure 2, the effect is stronger in the nonwords. Words and nonwords were named equally rapidly [383 vs. 386 msec; $F(1,35) = 1.83, n.s.$; $F(1,88) < 1.0$]. Some of the speed for nonwords may have occurred at the expense of the higher error rate. Responses were faster for items with initial fricatives than for those with final fricatives [345 vs. 424 msec; $F(1,35) = 223.16, p < .01$; $F(1,88) = 105.59, p < .01$]. The advantage for fricative initial items was greater in the second condition than in the first [66 vs. 92 msec; $F(1,35) = 15.66, p < .01$; $F(1,88) = 18.64, p < .01$]. Subjects were faster in the second condition than in the first [416 msec in the first, 354 msec in the second; $F(1,35) = 30.70, p < .01$; $F(1,88) = 434.30, p < .01$]. Changeability did not affect naming time (383 msec for the changeable items, 386 msec for the stable ones; both $F_s < 1$).

The results of Fox (1984) would suggest that some of the effect of the transition was obscured by rapid responses. Since subjects could begin an utterance without fully identifying the item, their rapid responses may be an incomplete measure of the naming task. Two further analyses were performed, one on all responses that began before the end of the item and one on all that began after the end of the item, in which all factors were treated as between factors. This allowed for the massively unequal cells that such a division entailed. While the degrees of freedom are inflated in this approach, the variance is inflated as well. The factors for subjects and tokens were, as expected, significant, showing that subjects and items have differences. As in the previous analysis, lexicality did not affect either the slow [$F(1,2184) < 1$] or the fast [$F(1,4181) < 1$] responses. Also, the items with initial fricatives elicited fast times, whether generally slow [$F(1,2184) = 336.31, p < .001$] or fast [$F(1,4181) = 286.55, p < .001$]. Changeability did not affect the fast times [$F(1,4181) < 1$], as had been found in the previous analysis. However, the slow responses showed an effect of changeability [$F(1,2184) = 4.62, p < .05$].

In the primary factor of interest, though, the prediction of an effect in the slow responses was borne out. The faster responses were unaffected by appropriateness of transitions [means of 263 msec for the matched and

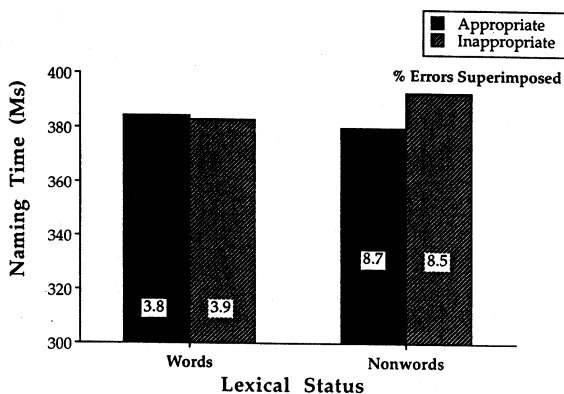


Figure 2. Effects of appropriateness of transition on auditory naming times for words and nonwords. Times have been adjusted by an average amount of time for the trigger to detect an onset for the particular item. The percentage of errors for each category is superimposed on each bar.

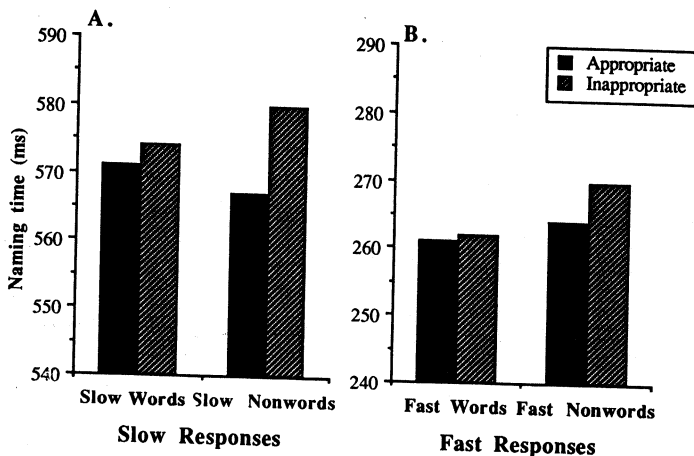


Figure 3. Effects of appropriateness of transition on auditory naming times for words and nonwords, separated into responses occurring after the end of the item (panel A) and those occurring before the end of the item (panel B). Times have been adjusted by an average amount of time for the trigger to detect an onset for the particular item.

266 msec for the mismatched; $F(1, 4181) < 1$], whereas the slower ones were [569 vs. 577 msec; $F(1, 2184) = 4.78, p < .05$]. (See Figure 3.) The interactions of appropriateness and lexicality were not significant and are shown in the figure only for comparison with the lexical decision results. No other interactions were significant. Although this test is liberal, it is supportive of the presence of an effect that is obscured by rapid responses.

Discussion

In the auditory naming task, the effect of subcategorical mismatches was lessened compared to the effect in the lexical decision task, showing up only in the slower responses. The naming times for words and nonwords were indistinguishable, though an effect of lexical status was evident in the greater error rate for nonwords than for words. As in Fox (1984), the effect of the lexicon appears in the responses that have not been launched before the lexical output could be used.

Recent work by Balota, Boland, and Shields (1989) has shown effects in production durations that were not evident in the onset latencies. It is quite possible that such effects might have been found with the present stimuli if the repetitions had been recorded. If conflicting information reaches the subject after the subject has already begun production, this might require some adjustment during the utterance, which would only appear in the duration measure. Fluent incorporation of previously unspecified phonemes has been shown to be possible for print (Whalen, 1990) and would presumably also be possible with auditory input. Such an analysis is not possible at this point.

The shadowing results support those of the lexical decision task. When the lexicon is involved in the response, phonetic mismatches affect both word and nonword judgments, whether the overriding cue comes before or after

the subordinate cue, and whether the change in fricative makes for a change in lexical status or not.

GENERAL DISCUSSION AND CONCLUSION

Subcategorical phonetic mismatches slow lexical decisions. In an auditory naming task, the same mismatches affect the onset time of both words and nonwords in cases that are slow enough for a lexical effect to appear. The presence of lexical involvement was further supported by a higher error rate for nonwords in the naming task. Despite a previously reported interaction between mismatches and lexical status (Streeter & Nigro, 1979), the present results suggest that phonetic mismatches are resolved before lexical access.

The discrepancy between these results and those of Streeter and Nigro (1979) may not be as severe as the exactly opposite conclusions suggest. Although Streeter and Nigro performed item analyses in some cases, they seem not to have done so for the lexical decision experiment. Thus the apparent difference in the strength of the lexicality/appropriateness interaction may not really exist, since it was significant for the subject analysis in the present work. In both studies a smaller difference was found for the matched and mismatched nonwords in the direction predicted by a phonetic account, but the interaction in the present study was not significant. Further refinements will be necessary to decide the issue more firmly.

The nature of the nonwords also seems to have an influence on the strength of the effect of mismatches. In the present results, there was an interaction of changeability (whether shifting from /s/ to /ʃ/ change the item from a word to a nonword or vice versa) with lexical status. The changeable nonwords were, by necessity, extremely like words. The stable nonwords, on the other

hand, tended to have few near neighbors (e.g., *giss* and *gish*). Lexical decisions on the stable nonwords were almost as fast as responses to words; only the changeable nonwords were noticeably slower. From the description in Streeter and Nigro (1979), it is not possible to tell how many of their nonwords had near neighbors. The two examples they do give (*faple* and *staces*) have near neighbors (*fable* and *stages*), and a large proportion of the others may have near neighbors too. Nonwords in the present study that did not change status received both fast judgments and low error rates. Having faster lexical decisions for nonwords may have made the difference in variance between words and nonwords less of a factor, allowing the phonetic mismatch effect to emerge.

If the resolution of phonetic mismatches is taking place prior to lexical access, the simplest assumption is that a level of phonemes is involved, though other possibilities exist. For example, it is possible that phonetic strings together with probabilities of accuracy are used to search the lexicon (as suggested by Ganong, 1980). In such a model, the resolution of the mismatches would not result in the categorical representation of the overriding cue's phoneme, but rather a graded version of that phoneme. However, if such a graded representation were used, it is difficult to see why the mismatch would have to be resolved before the gradation could occur. It seems more likely that the gradation would occur whether or not the cues were mismatched. Such an explanation would lead us to expect that the nonwords would show no effect, as described previously (Streeter & Nigro, 1979). The present results, though, show no interaction of the effect of mismatching with lexical status. Thus some level of abstraction prior to lexical access is suggested by these results.

If a level of phonemes is involved, there should be discontinuous use of the information in the acoustic signal, but some evidence has been presented as demonstrating continuous uptake. Warren and Marslen-Wilson (1987) used a gating paradigm to test whether subjects continuously access such partial information in the speech stream. They explicitly argue that such continuous uptake would deny a role to phonetic resolution of mismatches, which would imply a discrete phonemic output. However, the gating paradigm generates artificially complete utterances, which subjects must attempt to make sense of. That subjects do in fact tend to treat gated utterances as complete is demonstrated in a second study by Warren and Marslen-Wilson (1988), who examined a voicing distinction based on vowel duration. They found that, for the early gates (which are necessarily short), voiceless stop judgments predominated. Only when more duration was added was it possible for subjects to perceive voiced stops. Another aspect of the manner in which the information is taken up comprises effects of later information on earlier information. Changes in the vowel duration after the initial consonant can change the /b-/w/ boundary (Miller & Liberman, 1979). Whalen (1989) found, further, that when two distinctions (in this case, /æ/ vs. /e/ and /d/

vs. /t/) depended on the same vowel duration, the judgments interacted. Together, these results place a strong limitation on the interpretation of the gating studies. Although such studies give us insight into the kind and amount of information that is available, they do not necessarily show us just how the subjects use that information in ungated contexts. Thus it is still possible that coarticulatory information, which would be perceived separately in a gating study, would not normally generate a phonetic output until a larger portion of the signal had been processed.

Another approach to this kind of continuous uptake of phonetic information is put forth both in the phonetic refinement theory (Pisoni et al., 1985) and in the broad phonetic classification approach of Zue (1986; see also Huttenlocher & Zue, 1984). Despite their differences, both these systems assume that the broadest phonetic interpretation of the speech signal that nonetheless succeeds in retrieving a unique lexical entry will be the one used. In contrast, the present results show that even a cue that is going to fail, ultimately, to influence the phonetic interpretation is actively processed. It may well be that both the lexical decision task and the auditory naming task are still too close to arbitrary phonetic tasks for us to see the more normal workings of the lexicon. Certainly, if automatic speech recognition systems could be as successful as humans and yet only use the broad phonetic classifications, there would be a strong presumption that human listeners can and/or do use just such an approach. This theory, though, does not account for the accurate auditory naming of nonwords. The present results suggest instead that listeners make maximal use of the speech signal, even when it is counterproductive to do so.

REFERENCES

- BALOTA, D. A., BOLAND, J. E., & SHIELDS, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory & Language*, *28*, 14-36.
- BALOTA, D. A., & CHUMBLEY, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance*, *10*, 340-357.
- BALOTA, D. A., & CHUMBLEY, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory & Language*, *24*, 89-106.
- 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 & Performance*, *15*, 567-575.
- DARWIN, C. J. (1976). The perception of speech. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. VII. Language and speech* (pp. 175-226). New York: Academic Press.
- FOX, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception & Performance*, *10*, 526-540.
- GANONG, W. F., III. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception & Performance*, *6*, 110-125.
- GERRIG, R. J. (1986). Process and products of lexical access. *Language & Cognitive Processes*, *1*, 187-195.
- HUTTENLOCHER, D. P., & ZUE, V. W. (1984). A model of lexical access from partial phonetic information. *Proceedings of the Interna-*

- tional Conference on Acoustics, Speech and Signal Processing, 1984 (pp. 26.4.1-26.4.4.). New York: IEEE.
- KLATT, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day English*. Providence, RI: Brown University Press.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- MARSLÉN-WILSON, W. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125-150). Hillsdale, NJ: Erlbaum.
- MARSLÉN-WILSON, W. D., & WELSH, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- MARTIN, J. G., & BUNNELL, H. T. (1982). Perception of anticipatory coarticulation effects in vowel-stop consonant-vowel sequences. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 473-488.
- MCQUEEN, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception & Performance*, 17, 433-443.
- MILLER, J. L., & DEXTER, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 369-378.
- MILLER, J. L., & LIBERMAN, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semi-vowel. *Perception & Psychophysics*, 25, 457-465.
- MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, 76, 165-178.
- PISONI, D. B. (1978). Speech perception. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6, pp. 167-233). Hillsdale, NJ: Erlbaum.
- PISONI, D. B. (1985). Speech perception: Some new directions in research and theory. *Journal of the Acoustical Society of America*, 78, 381-388.
- PISONI, D. B., & LUCE, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, 25, 21-52.
- PISONI, D. B., NUSBAUM, H. C., LUCE, P. A., & SLOWIACZEK, L. M. (1985). Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, 4, 75-95.
- REMEZ, R. E., RUBIN, P. E., KATZ, M., & DODELSON, S. (1985, November). *On the influence of lexical status in phonetic perception*. Paper presented at the annual meeting of the Psychonomic Society, Boston.
- REPP, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92, 81-110.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- 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.
- TYLER, L. K., MARSLÉN-WILSON, W., RENTOUL, J., & HANNEY, P. (1988). Continuous and discontinuous access in spoken word-recognition: The role of derivational prefixes. *Journal of Memory & Language*, 27, 368-381.
- WARREN, P., & MARSLÉN-WILSON, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, 41, 262-275.
- WARREN, P., & MARSLÉN-WILSON, W. (1988). Cues to lexical choice: Discriminating place and voice. *Perception & Psychophysics*, 43, 21-30.
- WHALEN, D. H. (1982). *Perceptual effects of phonetic mismatches*. Unpublished doctoral dissertation, Yale University.
- WHALEN, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, 35, 49-64.
- WHALEN, D. H. (1989). Vowel and consonant judgments are not independent when cued by the same information. *Perception & Psychophysics*, 46, 284-292.
- WHALEN, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, 18, 3-35.
- WHALEN, D. H., & SAMUEL, A. G. (1985). Phonetic information is integrated across intervening nonlinguistic sounds. *Perception & Psychophysics*, 37, 579-587.
- WHALEN, D. H., WILEY, E. R., RUBIN, P. E., & COOPER, F. S. (1990). The Haskins Laboratories pulse code modulation (PCM) system. *Behavior Research Methods, Instruments, & Computers*, 22, 550-559.
- ZUE, V. W. (1986). Models of speech recognition: III. The role of analysis by synthesis in phonetic recognition. In P. Mermelstein (Ed.), *Proceedings of the Montreal Satellite Symposium on Speech Recognition (Twelfth International Congress on Acoustics)* (pp. 69-70).

NOTE

1. Pisoni and Luce (1987) have argued that what is normally studied is not lexical access but word recognition; but even recognizing that an item is a particular word is one "product" in Gerrig's (1986) sense of the access and differing levels of activation within the distributed representations of Seidenberg and McClelland's (1989) model. The amount of access that occurs is not relevant to the current study, which will therefore be described as dealing simply with "lexical access."

APPENDIX
Stimuli for Lexical Decision Task

		Initial s/f		Final s/f	
		Word	Nonword	Word	Nonword
s/f	s	1. soak (7)	shoak	goose (4)	goosh
change		2. sap (1)	shap	moss (9)	mosh
caused		3. soup (16)	shoup	bus (34)	buhsh
change in				[buss (1)]	
word/	4.	soap (22)	shoap	toss (9)	tosh
nonword	5.	soy (1)	shoy	fleece (0)	fleesh
status	6.	silk (12)	shilk	fuss (4)	fush
	f	1. shade (28)	sade	trash (2)	trass
		2. shaft (11)	saft	cash (36)	kass
		3. shout (9)	sout	gauche (1)	goass
		4. shut (46)	sut	bush (14)	boos
		5. shove (2)	suv	rash (1)	rass
		6. chef (9)	sef	wash (37)	woss
s/f		1. shoot (27)	shuke	mess (22)	giss
change		[shute (1)	chute (2)]		
caused no		suit (48)	suke	mesh (4)	gish
change in	2.	sift (0)	sipe	brass (19)	plless
word/		[sifted (3),	sifting (1)]		
nonword		shift (41)	shipe	brash (1)	pllesh
status	3.	sack (8)	sek	crass (2)	duss
		shack (1)	shek	crash (20)	dush
	4.	self (40)	sofe	lass (2)	koos
		shelf (12)	shofe	lash (6)	koosh
	5.	sock (4)	seeg	lease (10)	woas
		shock (31)	sheeg	leash (3)	woash
	6.	sake (41)	sud	douce (1)	froose
		shake (17)	shud	douche (0)	froosh

Note—Numbers in parentheses are word frequencies from Kučera and Francis (1967). Homonyms are listed in brackets.