Perceptual normalization for speaking rate: Effects of temporal distance

ROCHELLE S. NEWMAN and JAMES R. SAWUSCH
State University of New York, Buffalo, New York

A series of studies was undertaken to examine how rate normalization in speech perception would be influenced by the similarity, duration, and phonotactics of phonemes that were adjacent or distal from the initial, target phoneme. The duration of the adjacent (following) phoneme always had an effect on perception of the initial target. Neither phonotactics nor acoustic similarity seemed to have any influence on this rate normalization effect. However, effects of the duration of the nonadjacent (distal) phoneme were only found when that phoneme was temporally close to the target. These results suggest that there is a temporal window over which rate normalization occurs. In most cases, only the adjacent phoneme or adjacent two phonemes will fall within this window and thus influence perception of a phoneme distinction.

One of the fundamental issues in speech perception research involves the apparent lack of invariance between the acoustic signal and the listener's perception. Listeners somehow manage to perceive messages correctly, despite the variability in the acoustic signal caused by changes in speaking rate, talkers, and dialect. Researchers often have tried to examine each of these issues separately, in the hope that they would later be able to combine their findings into one theory.

One of the sources of variability in the acoustic signal is the rate at which a person speaks. People do not talk at a constant rate, and certain phonemes change substantially in duration as speaking rate changes (Crystal & House, 1982, 1990; Miller, Grosjean, & Lomanto, 1984; or see Miller, 1981, for a review of earlier work). In addition, talkers differ in their intrinsic rate of speech (see Crystal & House, 1988d), and some dialects either lengthen sounds or shorten them. The issue of rate change is especially important because some phonemic contrasts are cued, in whole or in part, by their duration. For instance, the /b/–/w/ manner contrast can be cued by differences in duration alone, with shorter initial transitions being heard as more "b-like" and longer transitions as more "w-like" (Liberman, Delattre, Gerstman, & Cooper, 1956; Miller & Liberman, 1979). However, when we listen to someone who talks very quickly, we still hear /w/ phonemes: they do not all sound like stops. Conversely, when we listen to someone who speaks very slowly, intended /b/s do not all sound like /w/s. Miller and Baer (1983) analyzed the transition

durations for /ba/ and /wa/ and found that for a given speaking rate, /w/ transitions were always longer than /b/ transitions. Across different speaking rates, however, the distributions overlapped for every speaker whom they examined. Obviously, we somehow compensate for the differences in speaking rate among talkers and among different tokens from any particular talker.

When researchers first began to examine how people compensated for changes in speaking rate, many focused on the immediately preceding words or syllables (see Miller, 1981, for a review). A few people examined more distal precursor effects (Kidd, 1989; Summerfield, 1981). From the work of Summerfield (1981) and Kidd (1989), it is clear that there are two components to the influence of a precursor phrase. One involves the influence of segments that immediately precede the target segment. The other involves the rate of speech of the phrase or sentence preceding the target. In addition to this work on the influence of preceding information upon the perception of duration-based phonetic contrasts, Miller and Liberman (1979) found that listeners take the apparent rate of speech *following* a particular phoneme into account when making duration-based distinctions. In their study, listeners identified tokens ranging from /ba/ to /wa/ with the syllable duration (vowel duration) varied between series. The listeners' category boundary between /b/ and /w/ changed with the variation in the following vowel duration. A longer vowel, suggesting a slower speaking rate, made the initial sound seem shorter in comparison. Transitions that might, at a faster rate, be too long for a /b/ would still be heard as /b/-like with this slower speech rate, and transitions would need to be especially long to be heard as /w/. This is seen in the data as more /b/ responses to the stimuli in the test series and a /b/–/w/ category boundary that occurs late in the series (near the /w/ end).

Likewise, a shorter vowel made the speaking rate seem faster and the beginning transitions seem longer, yielding more /w/ responses overall and a category boundary near

the /b/ end of the series. These results imply that a rate normalization process is taking place in speech perception. Phonetic contrasts that incorporate duration cues seem to be perceptually normalized, or scaled, to the rate of speech.

Miller and Liberman (1979) performed two additional experiments in order to explore this issue further. In one, they added a second syllable (/da/) to two of their original /ba/–/wa/ series. As in their first experiment, lengthening the vowel adjacent to the initial consonant shifted the category boundary to a longer transition duration. Lengthening the second syllable also had this effect, even though it was more removed, temporally, from the initial contrast. It is important to note that the difference in duration in the second syllable was entirely a difference in the duration of the vowel, since the /d/ transitions did not vary.

In their final experiment, Miller and Liberman (1979) examined the effect of adding a final consonant to the original consonant–vowel (CV) syllables. They added a final /d/ to two of the /ba/–/wa/ series and then compared the category boundaries for these new series with those for the original CV series. As before, lengthening the vowel led to a category boundary at a longer transition duration. Adding the final /d/ also had an effect, moving the category boundaries toward the /b/ end of the series. This is exactly *opposite* the effect of lengthening the vowel, which would lead to category boundary movement toward the /w/ end of the series.[1] Fowler (1983, 1987) suggests that the speech gesture for the consonant may overlay (and partially occlude) the longer gesture for the vowel. Or, perhaps the addition of the consonant provides a clear acoustic endpoint for the vowel. Either interpretation results in the vowel appearing to be shorter than it would without the consonant. This apparently shorter vowel would, in turn, make the initial segment seem longer, resulting in the category shift.

Much subsequent research has focused on explaining whether this rate normalization process is specific to speech or is part of a more general, auditory coding of sound (Diehl & Walsh, 1989; Pisoni, Carrell, and Gans, 1983). An alternative approach to describing the process is to map out the stimulus differences that influence rate normalization and examine how they relate to other linguistic coding. Returning to the /bad/–/wad/ stimuli of Miller and Liberman (1979), we know that the presence or absence of the final /d/ alters the listener's placement of the phonetic category boundary. Unfortunately, the authors did not examine the results further by systematically altering the length of this final consonant. There is no way of telling from these experiments whether systematic variation in the duration of nonvocalic segments can alter rate normalization. It is clear from their results that the *addition* of the final /d/ did shift the category boundary. However, the effect of changes *within* the final /d/ is a separate question.

Summerfield (1981) addressed this issue with the series /bi/–/pi/ and /biz/–/piz/. Altering vowel duration had the expected effect upon the /b/–/p/ boundary. A longer vowel produced more voiced "b" responses, moving the category boundary to longer voice onset times (VOTs). Conversely, a shorter vowel produced more voiceless "p" responses. Adding a final consonant shifted the locus of the category boundary, replicating the findings of Miller and Liberman (1979). Summerfield also altered the duration of the final consonant. This alteration had no effect, suggesting that the absolute duration of the final segment does not influence the initial distinction. However, it is unclear *why* the final segment had no effect. Perhaps it was too far removed temporally from the initial distinction. Perhaps consonants in general are not used in rate normalization, or perhaps they do not provide information strong enough to outweigh the effect of the vowel's duration. Perhaps a distal consonant could have an effect, but only if it was similar acoustically to the initial distinction.

Another approach to the issue of the nature of rate normalization processes is exemplified in the work of Miller and Dexter (1988). They found that when listeners responded faster to a stimulus, they did not show the same duration effects as when they responded more slowly. They used stimuli with an initial /b/–/p/ contrast and found the standard category boundary shift when manipulating the duration of the immediately following vowel (/i/), at least when they used a simple categorization task. However, when they altered the procedure to encourage listeners to respond faster, the results changed. When listeners responded slowly, they still showed the change in /b/–/p/ boundary characteristic of rate normalization. However, as listeners responded more quickly, their category boundaries converged. In the fastest responses, there was no change in the /b/–/p/ category boundary as a function of the duration of the following vowel. Presumably, listeners generated a response before the perceptual processing of the vowel was finished, so changes in vowel duration came too late to have an effect. This would explain why the change in vowel duration did not cause a boundary shift. However, the category boundary for the fastest responses was not located at the midpoint of the two boundaries for the slower responses, as one might expect. That is, the listeners' fastest responses did not treat all vowels as being of intermediate duration. Rather, their category boundary was similar to that found for shorter vowels in the slower responses.

Miller and Dexter (1988) suggested that, "although listeners do not always use all potentially relevant rate information within the syllable, they do use whatever rate information is available at the time of initiating the response. . . . Thus, listeners can base their decision on only the early portion of the syllable (leading to fast response times), but when doing so, they treat the syllable as if it were physically short" (p. 375). The listeners acted as if the vowel ended at the point in time at which they initiated the process of making a response. This interpretation suggests that the use of rate information is obligatory. A short-vowel effect was found for the fast responses, even though the listeners did not actually know what the vowel duration would be. Apparently they

could not avoid normalizing on the basis of the partial (and sometimes incorrect) data that they already had. This result also suggests that as information regarding rate is placed farther and farther from the initial contrast, there is less likelihood that listeners will process that information before making their decision. Thus, there should be a greater effect of speaking rate for phonemes closer to the durational contrast.

In all the studies examined so far, CV or CVC syllables were used, and in all of them effects of the vowel duration on perception of the previous, initial consonant were found. This leaves some questions unanswered, however. First, can a nonvocalic segment that varies in duration and follows the target influence perception of the target? Is the phonetic identity of the segment that follows the target a factor in rate normalization? Summerfield (1981) failed to find an effect with a consonant, but it was farther away from the initial distinction than was the vowel, which did show an effect. Thus, temporal distance of the contextual information from the target may be the critical factor. Second, is there an effect of acoustic (or phonetic) similarity? Are segments that share some features (either acoustic or phonetic) with the initial consonant more likely to serve as the basis for normalization than more dissimilar segments? Last, in what situations can nonadjacent segments affect rate normalization? Miller and Liberman (1979) found effects of a nonadjacent segment in one series (/bada/–/wada/), but no one has examined the temporal limits over which nonadjacent segments can demonstrate effects. We will discuss each of these issues in turn.

Phonetic Identity

Gay's (1978) studies of variation in segmental durations as a function of speaking rate show that the vowel appears to change more than do the surrounding stop consonants. The data on segmental durations of Crystal and House (1988a, 1988b, 1988c, 1988d) show that all phonetic segments do vary in duration with different speaking rates. However, vowel duration appears to be more variable than consonant duration (Crystal & House, 1988d). We might predict that there will be an effect of variation in phoneme duration only to the extent to which there is variation in production of that phoneme as a function of rate. That is, phonemes whose duration normally varies substantially with speaking rate might be expected to be more influential in rate normalization than phonemes that vary to a lesser extent. If this is the case, stop consonants should be less important for rate normalization than vowels. Furthermore, Miller and Baer (1983) found that with changes in speaking rate, the transitions for /w/ changed far more than those for /b/ (which remained close to constant). We might predict, then, that approximant (/w/ and /y/) duration also would affect rate normalization more than would stop consonant duration. Knowing that some phonemes are more strongly influenced by the effect of speaking rate, listeners may try to normalize the signal on the basis of those phonemes, re-

gardless of their position in the syllable. However, if listeners are responding quickly, they will only be able to use as much of the signal as they have already processed when they make their decision.

It is also possible that the vowel is particularly important for rate normalization. As the most sonorant phoneme in the syllable (or the sonority peak), vowels hold an important position in linguistic theory. Since every syllable has a sonority peak (usually, although not always, a vowel), but syllables need not have any other phonemes, the perceptual system might simply normalize on the basis of the vowel. Although we know of no theories that predict this explicitly, it would seem to be a plausible way to normalize for speaking rate. If this were the case, the vowel would serve as the basis of rate normalization, regardless of its position within the syllable, and other phonemes would have little or no effect. Thus, if the vocalic segments are the guideposts in rate normalization, we would expect nonadjacent vowels to still have a duration-based effect (except in speeded classification). Further, we would predict that variation in consonant duration should produce little or no rate normalization, regardless of the position of the consonant in the syllable. However, since phonemes seem to lie on a continuum from high to low sonority, perhaps the duration of more sonorant consonants, such as the approximants /w/ and /y/, would still have an effect.

Similarity

Another factor that may play a role in rate normalization is the acoustic similarity between the target and the later-occurring segment. For our purposes, we will assume that this acoustic similarity is based on segments having acoustic attributes, such as periodicity or aperiodicity, in common. The greater the number of common acoustic attributes (and the fewer the number of distinct acoustic attributes), the greater the similarity, and the greater the predicted rate normalization effect of the segment on the target.

A recent study by Green, Stevens, and Kuhl (1994) offers some tentative support for a principle of similarity. They compared rate normalization for four different types of CV syllables. The vowel identity was preserved across all four types, but the formant frequencies were shifted upward part way through the vowel to reflect a different talker (while the fundamental frequency remained unchanged), the fundamental was shifted (while the formant frequencies remained unchanged), neither was changed, or both were changed. Their results showed that when the fundamental was changed, only the initial vowel segment (the unaltered portion) influenced the listener's phonetic categorization. Although Green et al. did not find that changing the vowel formant frequencies altered the influence of vowel duration, subsequent work by Lotto, Kluender, and Holt (1994) has shown that a change in formant frequencies can produce results similar to the $F0$ results. That is, when either $F0$ or the formant frequencies change sufficiently, only the duration of the vowel

up to the change influences rate normalization. Thus, similarity (or continuity in F0 and/or the formants) appears to influence rate normalization.

It is also possible that phonetic similarity, in addition to acoustic similarity, will have an effect. That is, perhaps knowing that /p/ is very similar to /b/ would lead a listener to normalize one based on the other. If phonetic similarity were to have an effect, however, it might not be possible to separate this effect from one of acoustic similarity, since phonetic similarity is presumably based, at least in part, on similarity in the underlying acoustic information. Since our principle of similarity is based on this same acoustic information, distinguishing between these two types of similarity (phonetic and nonphonetic) may be difficult. In a similar vein, the similarity in articulatory gestures involved in the production of two phonemes may be involved in rate normalization. Again, separating this from acoustic (or phonetic) similarity would be difficult and is beyond the scope of this paper.

Adjacency

We expect that, all other things being equal, the closer in time that a particular acoustic segment occurs to the target segment, the greater the influence of that segment's duration on the perception of the target segment. The largest rate normalization effects should be found for acoustic segments adjacent to the target. This expectation is based on two lines of reasoning. First, as a consequence of coarticulation, adjacent segments carry mutually relevant information. This information would include speaking rate. Second, the on-line, real-time processing demands of spoken language understanding require that phoneme and word recognition keep pace with the signal. Consequently, remote information about speaking rate, especially when the remote information comes after the target phoneme, should be less likely to influence perception than adjacent information. However, if adjacency is the dominant principle in rate normalization, we would predict that adjacent segments, regardless of their phonetic identity, would show a strong effect. More distal phonemes should show a smaller effect, if at all, regardless of their similarity to the initial contrast or their sonority.

These principles all make different predictions about which phonemes in a sequence will have the greatest effect on an initial distinction. We will use the sequence /spoʃ/ ("spoash") as an example. If some notion of sonority is relevant for rate normalization, the vowel would have the strongest effect on perception of the initial consonant, even though it is separated from it by an intervening phoneme. If rate normalization is dependent on the degree to which phoneme classes tend to vary in production, we would similarly expect the vowel to have the strongest effect. Since stop consonants do not vary with speaking rate to the same degree as vowels vary, the influence of the /p/ would be smaller. On the other hand, the adjacency principle predicts that this intervening consonant would have the strongest effect, and that the effect of the vowel would be smaller, if present at all. If

rate normalization is based on acoustic similarity, the duration of the final /ʃ/ might well have a strong effect on perception of the initial contrast, since both phonemes are voiceless and aperiodic while the others are periodic and voiced. These different predictions were examined in the following experiments.

With these questions in mind, we decided to examine first the role of the similarity between the final consonant in CVC syllables and perception of a syllable-initial distinction. As mentioned above, both Miller and Liberman (1979) and Summerfield (1981) found that adding a final consonant to a CV syllable had an effect on perception of the initial consonant, but Summerfield found no effect of changing the duration of a final consonant. We wanted to follow up on these results, using a final consonant similar acoustically to the initial contrast.

EXPERIMENT 1

The purpose of this experiment was twofold: first, to determine whether a distal (not adjacent) consonant that follows the target phoneme will have the same type of effect on the perception of an initial duration-based contrast as an adjacent (following) vowel. Second, we wished to examine the role of acoustic similarity in rate normalization. For this study, we chose to use the /tʃ/–/ʃ/ distinction as the initial contrast. The affricate /tʃ/ is distinguished from the fricative /ʃ/ by the duration of fricative noise (Gerstman, 1957) and the "attack" or amplitude rise-time at onset (Cutting & Rosner, 1974). The /ʃ/ (as in "shoe") is typically longer in duration and has a more gradual attack, whereas the /tʃ/ (as in "chew") is shorter and has the abrupt attack characteristic of a stop. This contrast was chosen because duration is a potent cue to this phonetic distinction (see Repp, Liberman, Eccardt, & Pesetsky, 1978). Furthermore, Kluender and Walsh (1992) have demonstrated that frication duration alone is sufficient to signal the affricate/fricative distinction, and that this cue has a greater influence on perception than does rise-time variation.

The base syllables /tʃæs/ and /ʃæs/ were the starting point for the experiment. A synthetic /tʃæs/–/ʃæs/ series was generated in which the initial affricate–fricative distinction was cued by the duration of the initial, aperiodic segment of the syllable. From this series, two sets of stimuli were generated. For one set, three series were constructed with short-, intermediate-, and long-duration vowels. The duration of the final fricative /s/ was held constant. In the second set, the duration of the vowel was held constant. Here, the three series were constructed with short, intermediate, and long final /s/ durations.

An /s/ was chosen for the final consonant because it is aperiodic and because fricatives vary in duration as a function of speaking rate (Crystal & House, 1982, 1988c). The initial consonantal contrast and the final fricative are both aperiodic and voiceless, while the vowel is periodic and voiced. Thus, the consonants share an acoustic quality (aperiodicity) that distinguishes them from the vowel. If acoustic similarity influences the size of rate

normalization effects, the common aperiodicity of the initial contrast and the final /s/ should enhance any effect of the final consonant. That is, this similarity should make it more likely that the duration of the final consonant would affect perception of the initial contrast.

Method

Subjects. The listeners were 40 students from an introductory psychology course at the State University of New York at Buffalo who participated in the experiment for class credit. All listeners were native speakers of English and reported no history of a speech or hearing impairment. Eighteen participants listened to the three sets of stimuli which varied in the final /s/ duration, and 22 heard the sets with the variation in vowel duration. The data of 12 listeners were omitted from the analysis: 5 because of equipment failure, and 7 who could not classify the endpoints.[2] This left a total of 14 listeners in the final /s/ condition, and 14 in the vowel condition.

Stimuli. A male native speaker of English (D.J.R.) recorded the syllable /tʃæs/. This token was analyzed, and formant frequencies, over time, were computed using linear predictive coding. This information served as the starting point for synthesizing the syllable /tʃæs/ using the parallel mode of the cascade/parallel software synthesizer described by Klatt (1980). A complete description of the synthesis parameters for the syllable /tʃæs/ is provided in the Appendix. A 10-item series, ranging from /tʃ/ to /ʃ/, was created from this syllable by increasing the duration of the initial frication from 60 to 150 msec in 10-msec steps. Each successively longer stimulus was created by reduplicating the same 10-msec segment of synthesis parameters from the frication (time frames 15 and 20; see the Appendix). This resulted in an affricate–fricative series that varied only in the duration of the initial, steady-state portion of the frication. All other synthesis parameters were held constant across the series. Waveforms, showing the segmental durations of stimuli with 60- and 150-msec initial phonemes, are shown in Figure 1.

This series served as the base for all additional synthesis. The duration of the vowel /æ/ and final fricative /s/ in the base were used as the intermediate vowel and fricative durations. From this series, four additional series were created, one with a shorter /s/, one with a shorter vowel, one with a longer /s/, and one with a longer vowel. The longer /s/ series was created by duplicating nonadjacent 5-msec segments of the synthesis parameters for the frication. The shorter /s/ series was generated by deleting nonadjacent 5-msec

synthesis segments. The longer vowel series was created by duplicating nonadjacent 5-msec segments of the synthesis parameters, and the shorter series by deleting nonadjacent 5-msec segments. Care was taken in removing or reduplicating sections so as to maintain the approximate amplitude envelope of the original signal. The final durations for the /s/ and /æ/ portions are shown in Table 1. These durations are within the range of durations for vowels and fricatives reported by Crystal and House (1988b, 1988c, 1988d).

Creating these series resulted in 50 different syllables (a 10-item intermediate /æ/, intermediate /s/ series, two 10-item series with short and long /æ/, and two with short and long /s/). Two sets of stimuli were created, one with the /s/ varying (consisting of the short /s/ series, the intermediate series, and the long /s/ series), and the other with the vowel varying (consisting of the short /æ/ series, the intermediate series, and the long /æ/ series). Each set had a total of 30 stimuli, and participants listened to only one of the two sets.

Procedure. Listeners were run in small groups ranging from 1 to 5 at a time. Stimulus presentation and response collection were controlled by a Digital Equipment Corporation PDP-11/34 computer. The stimuli, which were stored on disk, were converted to analog form in real-time by a 12-bit, digital-to-analog converter at a 10-kHz sampling rate, low-pass filtered at 4.8 kHz, amplified, and presented binaurally through TDH-39 headphones. The syllables were presented in random order. Listeners were asked to rate the quality of the initial phoneme on a six-point scale, ranging from 1, *a good "ch,"* to 6, *a good "sh."* The use of ratings allows us to detect subtle differences within a category to which percentage data are not sensitive (Sawusch, 1976). After the stimulus had been presented, listeners responded by pressing the appropriate button on a computer-controlled response box. The presentation pace depended on the listeners' response speed. The next syllable was presented as soon as every listener had responded, or after an interval of 4.0 sec had elapsed, whichever came first. Responses from the first block of 60 trials were considered practice and were not included in subsequent data analysis. After the practice set, stimuli were presented in blocks of 90 trials (three repetitions of each of the 30 items), and all listeners received at least five blocks. Because of hardware problems on specific blocks, some listeners did not receive as many blocks as did others. In the varying vowel condition, 6 listeners received seven blocks (21 repetitions), 2 received six blocks (18 repetitions), and 6 received only five blocks (15 repetitions). In the varying /s/ series, 5 listeners received seven blocks (21 repetitions), 2 received six blocks (18 repetitions), and 7 received only five blocks (15 repetitions).
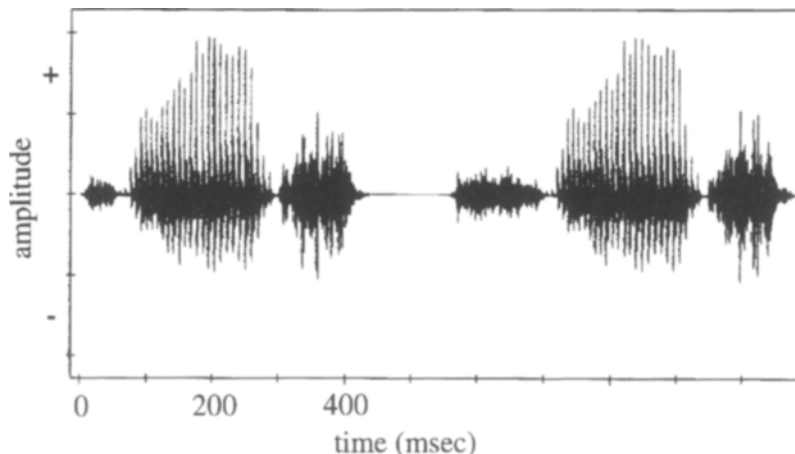

Figure 1. Waveforms of the end points of the /tʃæs/ and /ʃæs/ continuum, with /tʃæs/ (the shortest frication duration) on the left and /ʃæs/ (the longest frication duration) on the right.

Table 1
Vowel and Final Fricative Durations (in Milliseconds) in the /tʃæs/–/ʃæs/ Synthetic Speech Series of Experiment 1

Series	/æ/ Varying		/s/ Varying	
Short	/æ/	115	/æ/	230
	/s/	140	/s/	100
Intermediate	/æ/	230	/æ/	230
	/s/	140	/s/	140
Long	/æ/	395	/æ/	230
	/s/	140	/s/	190

Results and Discussion

For each listener, a mean rating was computed for each stimulus in each series. The /tʃ/–/ʃ/ category boundary for each listener was then determined for each series by linear interpolation between the rating responses for the two stimuli on either side of a neutral (3.5) response. We also tabulated the total percentages of "ch" responses given by each listener to all of the stimuli in each series.[3] While the movement of the category boundary should indicate changes in the perception of ambiguous stimuli, the overall percentage of "ch" responses to the series as a whole would include any changes away from the boundary of the series as well as those at the boundary. As noted by Samuel (1986) with respect to selective adaptation data, this overall percentage measure may be a more sensitive index of changes in perception. Two one-way analyses of variance (ANOVAs) with the factor of segment duration (short, intermediate, and long) were run on the category boundary data, one for the adjacent vowel duration data and one for the final fricative duration data. Two additional one-way ANOVAs, comparable to those run on the category boundary data, were run on the percentage data.

As expected, there was a significant effect of the variation in vowel duration on the affricate–fricative category boundary [$F(2,26) = 8.22, p < .01$]. Two of the subsequent paired comparisons were significant: The short and long vowel series were different [$t(13) = 3.61, p < .01$] and the intermediate and long vowel series were different [$t(13) = 2.75, p < .05$].[4] The difference between the short and intermediate vowel duration series was marginal [$t(13) = 1.89, p < .10$]. One interpretation of these results is that the longer vowel duration made the initial phonetic segment seem shorter in comparison. This resulted in a shift in the category boundary toward a longer initial segment duration, and thus more "ch" responses overall. The percentage data appeared to show slightly weaker results. There was a significant effect overall [$F(2,26) = 5.41, p < .05$], and the paired comparison between the short and long vowel series was significant [$t(13) = 2.65, p < .02$]. There was only a marginal difference between the intermediate and long series [$t(13) = 2.13, p < .06$], and no significant difference between the short and intermediate series [$t(13) = 1.76, p > .10$]. Table 2 shows the mean boundary locations and mean percentage "ch" responses (along with standard deviations) for the various conditions in this experiment. The rating functions for the three series that varied in vowel

duration, collapsed across the 14 listeners, are shown in the left-hand panel of Figure 2.

For the varying /s/ duration, there were no significant effects [$F < 1$ by category boundaries, and $F(2,26) = 1.12$, $p > .25$ by percentages]. Table 2 shows the mean boundary locations and percentage "ch" responses for the fricative duration conditions. The mean rating functions, collapsed across listeners, are shown in the right-hand panel of Figure 2. We were concerned that the vowel duration may have been so long as to prevent the /s/ from affecting the initial distinction. Since our goal was to determine whether it was possible for a distal consonant to affect the perception of the initial distinction, we wanted to make it as easy as possible for the consonant to have an effect. Therefore, we decided to rerun this condition with a shorter vowel duration, in order to make the final /s/ less temporally distant from the affricate–fricative distinction. At the same time, we made the short and long /s/ durations more extreme, with a short /s/ duration of 80 msec, the intermediate still 140 msec, and a long /s/ of 260 msec. Using these new stimuli, we ran 24 new listeners according to the procedure discussed previously. Of these, 4 had to be dropped from the experiment for a failure to classify the end points. The results for these new listeners were examined in the same way as those in the earlier condition, and with the same results. The duration of the final /s/ had no effect on the perception of the initial contrast [$F(2,38) < 1.0$].

These results replicate the earlier work which found an effect of variation in the duration of an adjacent vowel on an initial phonetic contrast (see Kluender & Walsh, 1988, for comparable results with the /tʃ/–/ʃ/ distinction). However, a later occurring consonant did not have this effect, even when its acoustic similarity to the initial distinction was putatively greater than that of the vowel to the initial distinction. If these results generalize to other cases, they may indicate that the major factor in subsequent information that influences rate normalization is the duration of the adjacent segment; or, they may suggest that the major factor is the duration of the vowel, since the current experiment cannot distinguish between these two possibilities.

These data do not provide support for a similarity principle, although they do not rule it out absolutely. It

Table 2
Experiment 1: /tʃæs/–/ʃæs/ Series

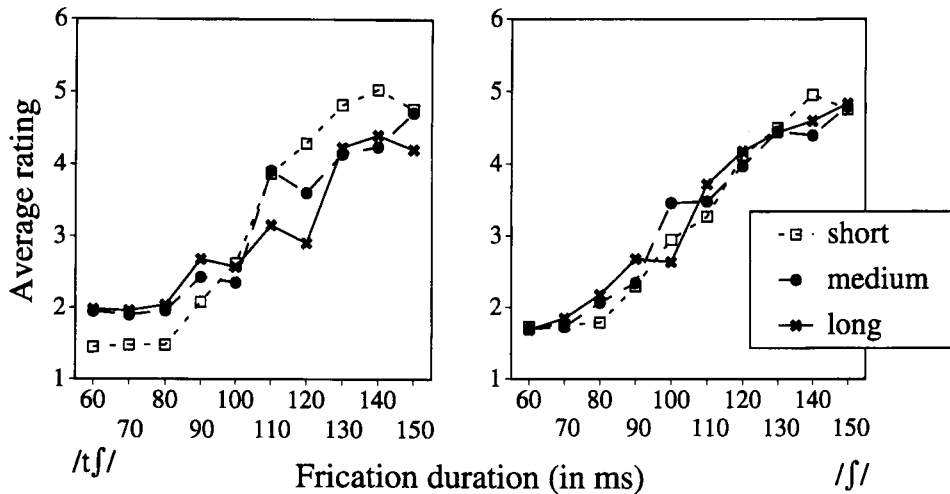|	Boundary Location (in msec)		Percentage "ch"	
	M	SD	M	SD
/æ/ vowel varying				
short series	107.2	8.48	56.03	8.76
intermediate series	114.8	14.16	60.26	7.43
long series	131.9	15.07	64.51	10.80
/s/ fricative varying				
short series	110.9	6.68	57.90	6.57
intermediate series	109.8	12.86	56.86	6.35
long series	110.0	9.69	55.89	6.63

Figure 2. Group rating functions for the /tʃæs/–/ʃæs/ series as a function of variation in vowel duration (left) and fricative duration (right).

is possible that among adjacent segments, a more similar one would receive a greater weighting. However, even if that were so, similarity would be a relatively weak stimulus factor. The type of similarity that we used here (shared periodicity) does influence selective adaptation (see Samuel, 1988). Even though the stimuli were constructed so that the nonadjacent segment /s/ was more similar to the initial phoneme (aperiodic, matching the initial /tʃ/–/ʃ/ contrast) while the adjacent vowel was less similar (periodic and not matching the initial /tʃ/–/ʃ/ contrast), no influence of the nonadjacent /s/ was found. This same pattern of no effect was found in a follow-up test with slightly different stimuli, so it seems to be a reliable result. Thus, at present, we have no evidence for a role of similarity with distal (nonadjacent) consonants in normalization for speaking rate.

EXPERIMENT 2

The results from Experiment 1 suggest that the durations of distal consonants have little or no effect on the perception of a syllable initial phoneme distinction. In comparison, the duration of an adjacent vowel has a robust influence. This leaves us with the question of, "Why the vowel?" Is it because it was a vowel per se, the nucleus of the syllable? Or is it because it was adjacent to the initial contrast? The former would suggest that attention is focused on the phonemes known to be most affected by speaking rate. The latter would suggest the possibility that sounds may always be normalized on the basis of the duration of other, adjacent, segments, regardless of what these segments are. In Experiment 2, we decided to move the vowel away from the initial contrast, turning it into a distal (nonadjacent) phoneme. In this experiment, the initial /tʃ/–/ʃ/ distinction was examined in the syllables /tʃwæs/ and /ʃwæs/. The duration of both the adjacent consonant (/w/) and the distal vowel (/æ/) was varied.

Method

Subjects. The listeners were 41 students from an introductory psychology course at the State University of New York at Buffalo who participated in the experiment for class credit. All listeners were native speakers of English and had no reported history of a speech or hearing impairment. Nineteen listeners heard the three series of stimuli that varied in the duration of the approximant /w/, and 22 heard the three series varying in /æ/ vowel duration. The data of 9 listeners were omitted from the analysis. In the approximant condition, 2 listeners could not categorize the stimuli, and 1 additional listener had to be dropped because of equipment failure. In the vowel condition, 3 listeners could not categorize the stimuli consistently, and another 3 were dropped because of equipment failure. This left a total of 16 listeners in the approximant condition, and 16 in the vowel condition.

Stimuli. A male, native speaker of English (J.R.S.) recorded the syllable /ʃwæs/ in the context of running speech. The stimulus was amplified, low-pass filtered at 4.8 kHz, digitized via a 12-bit, analog-to-digital converter at a 10-kHz sampling rate, and stored on computer disk. The syllable was excised from the carrier sentence, "Norton said _____ to me." The initial fricative segment from /ʃwæs/ was then separated from the remainder of the syllable, with the boundary being the zero-crossing preceding the first pitch pulse of the /w/. This particular speaker produced an unusually long fricative (172 msec), so 32 msec were cut from the beginning of the stimulus so that its length was more typical of that found in normal conversation. A 10-item continuum ranging from /ʃ/ to /tʃ/ was created by successively removing sections from the fricative onset to the zero crossing nearest 10 msec from onset. A linear amplitude ramp was used over the initial portion of each edited token. The duration of this ramp was varied with the duration of the frication, giving the stimuli a natural attack. This resulted in correlated durational and attack cues for the contrast, as would be found in natural speech. The resulting /ʃ/–/tʃ/ series ranged from 140 to 50 msec in duration in 10-msec steps, with the amplitude ramp varying from 54 to 1 msec in 9-msec steps. These duration values are similar to those of the synthetic series in Experiment 1, which ranged from 150 to 60 msec in duration. There was an additional 49 msec of coarticulation between the /ʃ/ and /w/. This section was not altered, and might be considered part of both phonemes.

The remainder of the syllable, /wæs/, was edited to create four new syllables: one with a shorter /w/, one with a shorter vowel, one with a longer /w/, and one with a longer vowel. The original /wæs/ served as the intermediate /w/ and intermediate /æ/. The durations

for the approximant /w/ and /æ/ vowel portions of these syllables in the different series are given in Table 3 (the values in the table do not include the 49 msec of coarticulation between the /ʃ/ and /w/ mentioned above).

As can be seen, the central, intermediate value is not halfway between the long and short values on a linear scale. Rather, it is centrally located on a logarithmic scale. That is, the proportional increase from intermediate to long is the same as the proportional increase from short to intermediate. These durations were chosen to reflect the distribution of durations for approximants and vowels in normal speech (see Crystal & House, 1988d).

We considered the first seven pitch pulses (after the /ʃ/–/w/ coarticulation) as belonging to the /w/ (rather than the /æ/) for editing purposes, because these pulses seemed to constitute the formant transitions (particularly the first formant) of the /w/. For the shorter /w/, the even-numbered pitch pulses were removed, for a total of 3 fewer pulses. For the longer /w/, each of the 7 pulses was duplicated, resulting in a stimulus with 14 pulses. In each case, a pitch pulse was identified as consisting of all of the waveform samples from the first sample after the last zero crossing before the pulse to the last sample before the last zero crossing before the next pulse. The /æ/ vowel contained 10 pitch pulses. For the short /æ/, every other pulse (the even-numbered pulses) was deleted, for a total of 5 fewer pulses. For the long /æ/, all 10 were duplicated. The short and long vowel-duration stimuli all contained the intermediate /w/ duration (69 msec), and the short and long /w/ stimuli had the intermediate vowel duration (104 msec).

These 5 tokens of /wæs/ were then spliced to the ends of each of the 10 members of the /tʃ/–/ʃ/ continuum, resulting in 50 different syllables. Two sets of stimuli were created, one with the /w/ varying (consisting of the short /w/ series, the intermediate series, and the long /w/ series), and the other with the vowel varying (consisting of the short /æ/ series, the intermediate series, and the long /æ/ series). Each set had a total of 30 stimuli, and listeners heard only one of the two sets.

Procedure. The procedure was almost identical to that used in Experiment 1. All participants listened to a practice block containing two occurrences of each of the 30 syllables in their set in random order. This was followed by blocks of 3 repetitions of each stimulus in the set. All listeners received at least five blocks. In the vowel condition, 3 listeners received six blocks, or 18 repetitions, whereas the remaining listeners all received five blocks, or 15 repetitions of each stimulus. In the approximant condition, 4 listeners received five blocks (15 repetitions), and the remaining 12 received six (18 repetitions) of each stimulus.

Results and Discussion

The basic data analysis procedure was the same as in Experiment 1. For each listener, a mean rating was computed for each stimulus in each series. The /tʃ/–/ʃ/ category boundary for each listener was then determined for each series, and a one-way ANOVA was run on these data for each of the two groups. As in Experiment 1, an ANOVA was also performed on the percentages of "ch" responses for each series.

As expected from the adjacency principle, the variation in duration of the approximant had a significant effect on the initial distinction [$F(2,30) = 4.27, p < .05$], based on the category boundaries. One of the subsequent paired comparisons was significant [short vs. long, $t(15) = 2.24, p < .05$] and another approached significance [intermediate vs. long, $t(15) = 2.08, p < .10$]. There was no significant difference between the short and intermediate duration series [$t(15) = 0.11, p > .10$]. With the percentage data, the overall effect was significant [$F(2,30) = 8.58, p < .005$], as were two of the paired comparisons [short vs. long, $t(15) = 2.93, p < .02$, and intermediate vs. long, $t(15) = 3.36, p < .01$]. There was still no significant difference between the effects of the short and intermediate /w/ duration series [$t(15) = 0.28, p > .10$]. Table 4 shows the mean boundary locations and mean percentages (along with standard deviations) for the various conditions in this experiment. The left panel of Figure 3 shows the rating functions for these three series, averaged across the listeners. (The values of the frication duration on the abscissa do include the portion coarticulated with the /w/.)

There was no effect of vowel duration on the initial /tʃ/–/ʃ/ distinction [$F(2,30) < 1.0$ for both category boundary and percentage data]. The data are presented in Table 4 and the three rating functions are shown in the right panel of Figure 3. The results from this experiment, together with those of Experiment 1, are consistent with an adjacency principle in rate normalization. The adjacent segment always had an effect, whether it was a vowel or not. The vowel only had an effect when it was adjacent.

We should also consider one alternative explanation for our results. It is possible that the /w/ had an effect because, like vowels, it is strongly influenced by speaking rate (Crystal & House, 1988d; Miller & Baer, 1983). If we pay attention to vowel duration because it carries so much relevant information, it would not be surprising if we also attended to approximant duration. Given this line of reasoning, replacing the /w/ with a stop consonant should reduce or eliminate the effects of the adjacent consonant, since stop consonants are less influenced by speaking rate than are approximants (Miller & Baer, 1983).

EXPERIMENT 3

Experiment 3 was very similar to Experiment 2, with the exception that the stop consonant /k/ was used instead of the approximant /w/ as the adjacent consonant, and a different vowel was used. The syllables in this experiment were /tʃkas/ and /ʃkas/. If changing the duration of the /k/ affects the perception of the initial contrast, we can state with some confidence that any adjacent phoneme can affect the perception of the initial duration distinction. If the stop consonant does not have an effect, however, we must assume that the results from Experiment 2 were the result of using an approximant as the consonant, lending support to a vocalic principle.

Four general outcomes are possible from this experiment: (1) the variation in neither the /k/ duration nor the

Table 3
Glide and Vowel Durations (in Milliseconds) in the /tʃwæs/–/ʃwæs/ Natural Speech Series of Experiment 2

Series	/w/ Varying		/æ/ Varying	
Short	/w/	41	/w/	69
	/æ/	104	/æ/	52
Intermediate	/w/	69	/w/	69
	/æ/	104	/æ/	104
Long	/w/	132	/w/	69
	/æ/	104	/æ/	197

Table 4
Experiment 2: /tʃwæs/–/ʃwæs/ Series

	Boundary Location (in msec)		Percentage "ch"	
	M	SD	M	SD
/w/ glide varying				
short series	138.0	6.26	50.10	5.39
intermediate series	137.9	6.30	49.89	5.31
long series	141.4	8.55	54.33	7.46
/æ/ vowel varying				
short series	137.7	9.53	51.53	9.34
intermediate series	136.9	11.50	51.13	9.67
long series	137.0	9.26	51.47	8.85

/a/ duration has an effect; (2) the effect of the duration of the adjacent consonant dominates perception; (3) the effect of the vowel dominates perception; or (4) the durations of both phonemes contribute to perception. The first possibility would suggest that only segments that are both vocalic and adjacent can have duration-based effects. If the /k/ duration has an effect, while the vowel duration either has no effect or a much weaker one, the results from this experiment alone would support a model of rate normalization based on a principle of adjacency. It seems unlikely that the nonadjacent vowel alone would have an effect, since it did not have one in the previous experiment. However, were this to occur, we would have to assume that the most nearly adjacent segment that normally varies with speaking rate tends to have an effect. Since approximants can be influenced by speaking rate, the /w/ had an effect in the prior experiment. Since stop consonants are less influenced by speaking rate, normalization waits until the next phoneme, the vowel, occurs. Although this explanation is logical, it seems rather inelegant. Finally, if the /k/ and /a/ both have an effect, it would suggest that there may be both an effect of phoneme identity and an adjacency effect. Again, it would

be difficult to explain why the nonadjacent vowel had an effect here but not in the previous experiment.

Method

Subjects. The listeners were 33 volunteers from the State University of New York at Buffalo and the surrounding community. They were paid $5 for their participation in the experiment. All were native speakers of English with no reported history of a speech or hearing disorder. Fourteen listeners heard the stimuli with the varying /k/ duration, and 19 heard the stimuli with the varying /a/ vowel duration. Of these, 9 had to be dropped from the analysis of the data because they could not consistently classify the end points. This left 11 listeners in the condition with the varying /k/ duration, and 13 who heard the stimuli with the varying /a/ duration.

Stimuli. A male native speaker of English (J.R.S.) recorded the syllable /ʃkas/ in the context of running speech. His speech was amplified, low-pass filtered at 4.8 kHz, digitized via a 12-bit, analog-to-digital converter at a 10-kHz sampling rate, and stored on computer disk. The syllable was excised from the carrier sentence, "Norton said _____ to me." The initial /ʃ/ fricative segment from /ʃkas/ was then spliced from the remainder of the syllable at the beginning of the closure for the /k/. A 10-item continuum ranging from /ʃ/ to /tʃ/ was created by removing successive 10-msec sections from the /ʃ/ fricative onset, as previously described in Experiment 2. A linear amplitude ramp was used over the initial portion of each edited token. The duration of this ramp was varied with the duration of the frication. This gave the stimuli a more natural attack and resulted in correlated durational and attack cues for the contrast, as would be found in natural speech. The resulting /ʃ/–/tʃ/ series ranged from 150 to 60 msec in duration in 10-msec steps, with the amplitude ramp varying from 60 to 6 msec in 9-msec steps.

The remainder of the syllable, /kas/, was edited to create four new syllables: one with a shorter /k/, one with a shorter vowel, one with a longer /k/, and one with a longer vowel. The approximate durations for the consonant and vowel portions of these syllables are given in Table 5.[5] The vowel and stop durations are within the range reported by Crystal and House (1988a, 1988b, 1988d) for American English.

We considered the first 4 pitch pulses as belonging to the /k/ (rather than the /a/) for editing purposes, because these seemed to correspond to the formant transitions (particularly the first formant transition). For the shorter /k/, 2 of these pulses were deleted (the
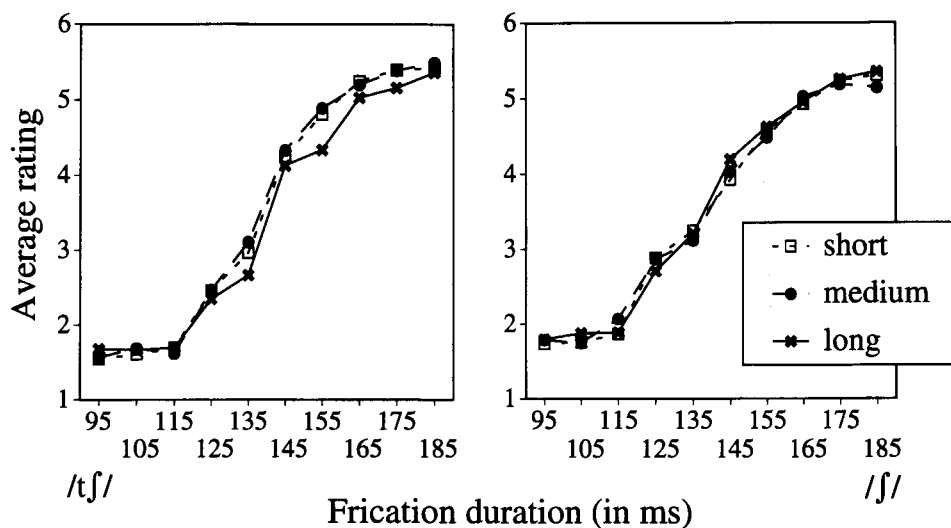

Figure 3. Group rating functions for the /tʃwæs/–/ʃwæs/ series as a function of variation in semivowel duration (left) and vowel duration (right).

even-numbered pulses), as well as 16.5 msec of the burst and aspiration (the original burst and aspiration together were about 31 msec in duration). Short (less than 3.5-msec duration) nonadjacent sections of the burst and aspiration were deleted so that the overall amplitude profile remained the same and any frequency changes in the spectrum were preserved. For the longer /k/, each of the 4 pitch pulses was duplicated, and a total of 15.5 msec of burst plus aspiration was also duplicated. In order to prevent the perception of frozen noise, only very short sections of the burst and aspiration were duplicated (no more than 3.5 msec), and these portions were spread throughout the aperiodic portion of the /k/. The closure duration between the initial fricative/affricate and the /k/ was left at its original duration of 69.1 msec in all stimuli (the closure is not included in the segment durations in Table 5). The closure duration was not modified, because in the data of Crystal and House (1988a, 1988d) there was only a 4-msec difference in average closure duration between fast and slow talkers. The short vowel token was created by deleting 8 pitch pulses. Beginning with the 6th pulse after the onset of voicing, every other pulse was deleted. The long vowel token was created by duplicating every pulse from the 6th to the next to last, or 15 pulses. The deletion and duplication of pitch pulses was done as previously described in Experiment 2.

The short and long vowel durations both occurred with the intermediate stop consonant duration, and the short and long /k/ durations occurred with the intermediate vowel duration. These five tokens of /kas/ were then spliced to the ends of each of the 10 members of the /tʃ/–/ʃ/ continuum, resulting in 50 different syllables. Two sets of stimuli were created, one with the /k/ varying (consisting of the short /k/ series, the intermediate series, and the long /k/ series), and the other with the vowel varying (consisting of the short /a/ series, the intermediate series, and the long /a/ series). Each set had a total of 30 stimuli, and listeners heard only one of the two sets.

Procedure. The procedure was identical to that used in Experiments 1 and 2. All listeners heard a practice block containing two occurrences of each of the 30 syllables in their set in random order. This was followed by blocks of three repetitions of each stimulus. Most listeners heard seven blocks (21 repetitions of each stimulus). However, because of hardware problems on specific blocks, 2 listeners in the varying vowel condition only heard six blocks (or 18 repetitions).

Results and Discussion

For each listener, an average (mean) rating was computed for each stimulus in each series. The /tʃ/–/ʃ/ category boundary for each listener was then determined for each series, and a one-way ANOVA was run on these data for each of the two groups. As in Experiments 1 and 2, an ANOVA was also performed on the percentages of "ch" responses for each series.

For the stimuli varying in /k/ duration, the results showed a significant effect of the consonant duration on the affricate–fricative category boundary [$F(2,20) = 3.86$, $p < .05$]. The subsequent paired comparison of the short

and long series was significant [$t(10) = 2.66$, $p < .05$], but those on the short and intermediate series and the intermediate and long series were not significant [$t(10) = 1.61$, $p > .10$, and $t(10) = 1.07$, $p > .10$, respectively]. The percentage data also showed a significant overall effect [$F(2,10) = 5.77$, $p < .01$], as well as significant pairwise comparisons between the intermediate and long series [$t(10) = 4.12$, $p < .01$] and the short and long series [$t(10) = 3.31$, $p < .01$]. The difference between the short and intermediate series was not significant [$t(10) = 1.33$, $p > .10$]. The effect of a longer /k/ duration was as if it made the initial /tʃ/–/ʃ/ segments seem shorter in comparison. This resulted in an overall shift in the category boundary toward longer initial segment durations and more "ch" responses overall as the /k/ duration was increased. Table 6 shows the mean boundary locations and mean percentages (along with standard deviations) for the various conditions in this experiment. The mean rating functions (across listeners) for the three series with the /k/ duration varied are shown in the left panel of Figure 4.

For the stimuli varying in vowel duration, there were no significant changes in the category boundary [$F(2,24) < 1.0$]. The results for the percentage of "ch" responses to each series showed a similar lack of any effect of vowel duration [$F(2,24) < 1.0$]. The vowel series rating data are shown in the right-hand panel of Figure 4.

The results from the varying /k/ series demonstrate that an adjacent stop consonant can have an effect on an initial duration-based contrast. This is consistent with an adjacency principle of rate normalization, and it suggests that the effects found for the /w/ in the previous experiment were not simply due to the vocalic nature of the consonant. The lack of results from the vowel also provide support for an adjacency model.

So far, however, we have only used the "sh"–"ch" series. Perhaps more importantly, both ends of the series we used in Experiment 3, as well as the /tʃ/ end of the series from Experiment 2, are phonotactically illegal in English, and it is possible that our effects might not generalize to legal series. Some researchers have suggested that rate normalization is a general auditory process (Diehl & Walsh, 1989; Pisoni et al., 1983). In this case, we would not expect that our results reflect an influence of phonotactics. Nevertheless, we still believed it important to ascertain whether phonotactic legality might influence the normalization process. With this in mind, we decided to see whether our results would generalize to a series that was phonotactically legal in English.

EXPERIMENT 4

In this experiment, we created a /twæs/–/swæs/ series that contained the same /wæs/ portion as the /tʃwæs/–/ʃwæs/ series in Experiment 2. Basically, we removed the /tʃ/–/ʃ/ portion from the Experiment 2 stimuli and replaced it with frication designed to cue a /t/–/s/ distinction. This resulted in a phonotactically legal series with the same acoustic properties and variation in duration of adjacent and remote phonemes as those of the phonotac-

Table 5
Stop Consonant and Vowel Durations (in Milliseconds) in the /tʃkas/–/ʃkas/ Natural Speech Stimuli Used in Experiment 3

Series	/k/ Varying		/a/ Varying	
Short	/k/	28	/k/	60
	/a/	159	/a/	82
Intermediate	/k/	60	/k/	60
	/a/	159	/a/	159
Long	/k/	104	/k/	60
	/a/	159	/a/	293

Table 6
Experiment 3: /tʃkas/–/ʃkas/

	Boundary Location (in msec)		Percentage "ch"	
	M	SD	M	SD
/k/ stop varying				
short series	88.1	8.91	42.74	7.58
intermediate series	93.0	11.50	46.04	8.50
long series	95.1	9.23	49.02	7.35
/a/ vowel varying				
short series	91.3	12.18	42.61	11.58
intermediate series	92.6	8.49	43.95	8.17
long series	93.1	10.99	43.59	10.22

tically illegal series in Experiment 2. Furthermore, we created this series so that the /t/ end was similar in duration to the /tʃ/ end of the /ʃwæs/–/tʃwæs/ series and the /s/ was similar to the /ʃ/ end in duration. Thus, the stimuli in this experiment were phonotactically legal but similar to those of Experiment 2 in terms of duration. If the results from this experiment mirror those in Experiment 2, (that is, if the approximant duration has an effect, but the vowel duration does not) it would suggest that legality per se does not influence rate normalization, at least not in this context. Furthermore, it would suggest that the effect is not limited to the /tʃ/–/ʃ/ series, and it would provide further support for an adjacency principle in rate normalization. If, instead, both the approximant and vowel durations produce an effect, it would strongly suggest that rate normalization is influenced by the phonotactics of a listener's language.

Method

Subjects. The listeners were 41 students from an introductory psychology course at the State University of New York at Buffalo who participated in the experiment for class credit. All listeners were native speakers of English and had no reported history of a speech or hearing impairment. Twenty listeners heard the three sets of stimuli that varied in /w/ duration, and 21 heard the sets with variation in vowel duration. Two listeners in the vowel series were dropped for a failure to categorize one or more end points accurately, and 1 additional listener was dropped because of an equipment failure. This left a total of 19 listeners in the approximant condition, and 19 in the vowel condition.

Stimuli. A male native speaker of English (J.R.S.) recorded the syllables /swæs/ and /twæs/ in the context of running speech. The stimuli were amplified, low-pass filtered at 4.8 kHz, digitized via a 12-bit, analog-to-digital converter at a 10-kHz sampling rate, and stored on computer disk. The syllables were excised from the carrier sentence "Norton said _____ to me." The initial consonant segments were then separated from the remainder of the syllables, with the boundary being the start of periodicity. In order to make a series ranging from /t/ to /s/, we removed successive sections of approximately 10 msec from the onset of the /s/, as previously described. Ainsworth (1977) has previously shown that variation in frication duration is sufficient to cue a /t/–/s/ distinction. Unfortunately, the resulting /t/ to /s/ series included 15 members. We did not want to run the experiment with this many stimuli, but also did not want any loss of sensitivity to category boundary movement. With this in mind, we pilot tested the base series to determine the location of the category boundary. The listeners were very consistent, placing their category boundary between the 7th and 8th members of the continuum. We decided to keep the 1st, 3rd, 5th, 6th, 7th, 8th, 9th, 11th, 13th, and 15th members of the continuum, deleting alternating members toward both ends of the continuum. This left us with a 10-member series with initial frication durations of 42, 62, 82, 92, 102, 112, 122, 140, 161, and 178 msec. The duration of the most extreme /s/ was 178 msec, a bit longer than the longest member of the /tʃ/–/ʃ/ continuum of Experiment 2 (152 msec). The /t/ end was a bit shorter than was the /tʃ/ end of the earlier series (/t/ was 42 msec, /tʃ/ was 62 msec). Thus, we had a slightly greater duration range in this experiment than in the /tʃ/–/ʃ/ experiments.

We next took the five /wæs/ syllables from Experiment 2 and appended them to our 10-member /t/–/s/ items. (We first deleted 49 msec of the /ʃ/–/w/ coarticulation that would have caused the stimuli to sound unnatural.) This resulted in five /twæs/–/swæs/ series (50 syllables) whose /wæs/ portions were nearly identical to the analogous /tʃwæs/–/ʃwæs/ series (Experiment 2). Two sets of stimuli were created, one with the /w/ varying (consisting of the
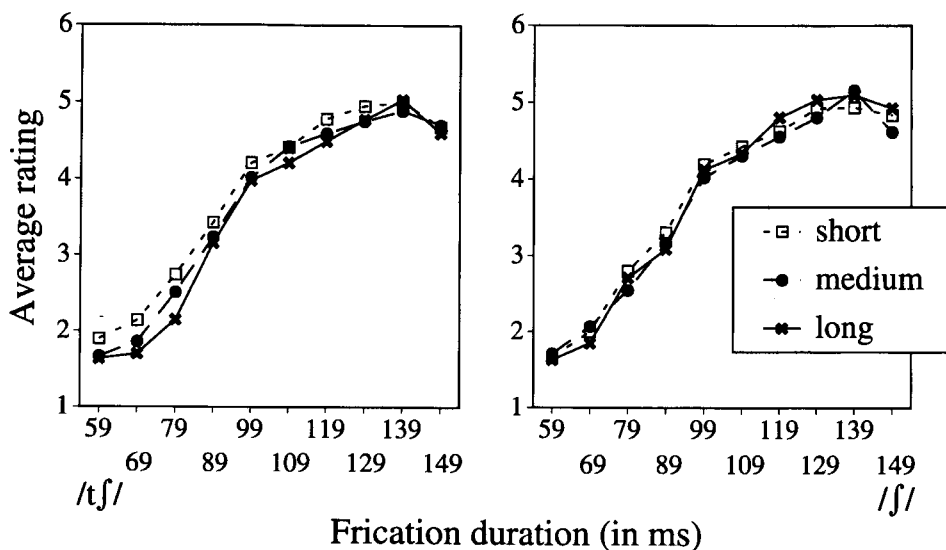

Figure 4. Group rating functions for the /tʃkas/–/ʃkas/ series as a function of variation in stop consonant duration (left) and vowel duration (right).

short /w/ series, the intermediate series, and the long /w/ series), and the other with the vowel varying (consisting of the short /æ/ series, the intermediate series, and the long /æ/ series). Each set had a total of 30 stimuli, and listeners heard only one of the two sets. The durations of the /w/ and /æ/ portions of the stimuli have been given previously, in Table 2.

Procedure. The procedure was identical to that used in the previous experiments. All listeners heard a practice block containing two occurrences of each of the 30 syllables in their set in random order. This was followed by five blocks of three repetitions (15 presentations) of each stimulus.

Results and Discussion

For each listener, an average (mean) rating was computed for each stimulus in each series. The /t/–/s/ category boundary for each listener was then determined for each series, and a one-way ANOVA was run on these data for each of the two groups. As in the previous experiments, an ANOVA was also performed on the percentages of "t" responses for each series.

As expected, the duration of the approximant had a significant effect on the labeling of the initial distinction [$F(2,36) = 10.93, p < .0001$]. Two of the subsequent paired comparisons were significant: short versus long [$t(18) = 3.86, p < .005$] and intermediate versus long [$t(18) = 3.22, p < .005$]; the short versus intermediate was not significant [$t(18) = 1.50, p > .10$]. With the percentage data, the results were similar. The overall effect was significant [$F(2,36) = 9.90, p < .0001$], as were all three of the paired comparisons: short versus intermediate [$t(18) = 2.35, p < .05$], intermediate versus long [$t(18) = 2.69, p < .01$], and short versus long [$t(18) = 3.65, p < .005$]. Table 7 shows the mean boundary locations and mean percentages (along with standard deviations) for the various conditions in this experiment. The rating functions for the three /w/ varying series are shown in the left panel of Figure 5; the right panel shows the results for the three vowel duration series.

The vowel results were not as clear cut. Using the category boundaries, there was no effect of vowel duration on the initial distinction [$F(2,36) = 1.05, p > .10$]. However, the percentage data did show an effect [$F(2,36) = 8.30, p < .005$]. Paired comparisons on the percentage data showed a significant effect of vowel duration for the long versus intermediate comparison [$t(18) = 3.96, p < .001$] and for the short versus long comparison [$t(18) = 3.26, p < .005$], but no effect for the short versus intermediate condition [$t(18) = 1.43, p > .10$]. This suggests that whatever effect the vowel duration was having on the /t/–/s/ series, it was not at the category boundary.

In order to examine this further, we did a 3 × 10 ANOVA (with factors of vowel duration and series stimulus) with follow-up tests to examine the effect of vowel duration on each item in the /t/–/s/ series. We hoped this would allow us to determine precisely where the effect was coming from, since it was not at the boundary. There was a main effect of the duration of the vowel [$F(2,36) = 5.95, p < .01$] and a main effect of stimulus in the test series [$F(9,164) = 238.3, p < .001$], but the interaction was marginal [$F(18,324) = 1.53, p < .10$]. Because there

were 10 simple effects tests, we decided to adopt a more conservative procedure than the standard t test. We used the Greenhouse-Geisser adjusted degrees of freedom. There was no effect of vowel duration in the first stimulus (the /t/ end, $F < 1.0$). The effects on the second and third stimuli in the series were both marginal ($F = 3.29$, $p < .10$ and $F = 3.07, p < .10$). The fourth member of the series showed a significant effect of vowel duration ($F = 5.12, p < .05$), whereas the effect at the fifth stimulus was marginal ($F = 2.94, p < .10$). There was no significant effect on the sixth stimulus ($F < 1.0$), and there was a marginal effect of vowel duration on the seventh stimulus ($F = 3.29, p < .10$). There was no significant effect of vowel duration for the last three members (the /s/ end) of the test series ($F < 1.0, F < 1.0$, and $F = 1.22$, respectively).

This asymmetric result may seem rather surprising. One possible explanation is that the /t/ end of the series was more ambiguous than the /s/ end. Since our series were made from truncating the /s/, it would not be too surprising if the /t/ end was less clear. Because ambiguous stimuli may be more susceptible to other influences, this could potentially explain these results. One way to examine this is to compare listeners' ratings for the end points of the series. If the /t/ stimuli are more ambiguous overall, the /t/ end point should receive lower ratings (more neutral) than the /s/ end point. We performed a 2 (phoneme varying: vowel varying or approximant varying) × 3 (series: short, medium, or long) × 2 (end point: t or s) ANOVA on the rating data for the end points. We found a marginal effect of end point ($F = 3.413, p < .08$), but no other significant or marginal effects. The marginal effect of end point, however, was caused by the /t/ end having *better* (more extreme) ratings than the /s/ end point, not the reverse. That is, the /t/ end point was rated as being as good as or better than the /s/ end point. This suggests that the /t/ end of the series was unlikely to be more ambiguous than the /s/ end.

The second possibility is based on the fact that the /t/ end is the shorter end of the series. Perhaps the effect was found here because rate normalization operates only over a limited temporal portion of the waveform. If only the portion of the stimulus within a limited temporal window around the target influences the perception of the target, and all or most of the vowel was included before the end of this temporal window on the shortest

Table 7
Experiment 4: /twæs/–/swæs/

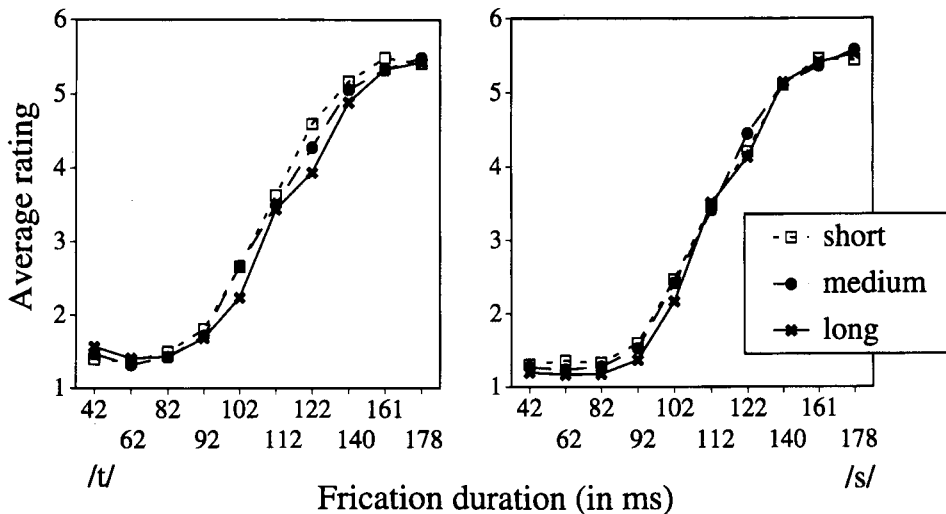|	Boundary Location (in msec)		Percentage "t"	
	M	*SD*	*M*	*SD*
/w/ glide varying				
short series	101.4	7.72	55.56	7.56
intermediate series	102.4	8.01	57.22	7.20
long series	106.8	10.17	60.16	8.86
/æ/ vowel varying				
short series	102.8	11.11	57.45	9.63
intermediate series	102.2	8.74	58.38	8.71
long series	103.9	10.24	60.00	9.04

Figure 5. Group rating functions for the /twæs/–/swæs/ series as a function of variation in semivowel duration (left) and vowel duration (right).

members of this series, then some effect might be observed for shorter stimuli. Thus, a modification of the adjacency principle to incorporate a moving temporal window for rate normalization would appear to be consistent with the data from this experiment.

Table 8 lists the duration from the beginning of the initial contrast (the beginning of the target phoneme) to the end of the segment whose duration was varied between series in each of the first four experiments. The first two columns list the durations to the end of the adjacent segment (/æ/ in Experiment 1, /w/ in Experiments 2 and 4, /k/ in Experiment 3, and /l/ in Experiment 5), and the second two columns list the durations to the end of the nonadjacent segment (/s/ for Experiment 1, /æ/ for Experiments 2 and 4, /a/ for Experiment 3, and /o/ for Experiment 5). For both the adjacent and the nonadjacent segments, the durations are listed for the longest and shortest durations of these segments. For greater clarity, we have not listed the durations for the intermediate versions.

The duration from the start of the /twæs/ syllable to the end of the vowel is relatively short, even for the longest vowel (317 msec, on the far right). Perhaps the temporal window was long enough to include part of the vowel in this series. But this leaves the question as to why we did not see a similar effect at the /tʃ/ end of the /tʃwæs/–/ʃwæs/ series, since it was only slightly longer in length (376 msec; see Table 8). One possibility is that the slight length difference was just enough to make the effect too small to see in our overall analysis. With this in mind, we went back and did an items analysis on these data analogous to the one we did for the /twæs/–/swæs/ series. We found no significant effects, and the only marginal effects were at the long end of the series, not the short end. [For the 1st item, $F < 1.0$; for the 2nd and 3rd items, there was a marginal effect, with $F(2,14) = 3.10, p < .10$, and $F(2,14) = 3.19, p < .10$, respectively. All the rest of the items were not significant: $F = 1.46$ for the 4th; $F < 1.0$ for the 5th; $F = 1.10$ for the 6th; for the 7th, 8th, and 9th

items, $F < 1.0$; for the 10th, $F = 1.08$.] This suggests that there simply was no effect rather than a slight effect at the /tʃ/ end that our overall tests did not detect.

A second possibility is that the size of the rate normalization temporal window for information following the target is not fixed. Rather, it varies with the quality of the acoustic-phonetic information that the listener receives. As acoustic-phonetic information accumulates for each segment, more and more information about speaking rate is also processed. If the acoustic-phonetic information is of relatively high quality (natural speech, spoken in a noise-free environment), then segmental decisions are made relatively rapidly, and little effect of a remote segment is observed in rate normalization. Conversely, if the acoustic-phonetic information is of lower

Table 8
Duration (in Milliseconds) From Onset of Syllable
to End of Altered Segment

Experiment	Target End Point	Segment			
		Adjacent		Nonadjacent	
		Short	Long	Short	Long
Experiment 1					
/tʃæs/–/ʃæs/	"ch"	175	455	390	480
	"sh"	265	545	480	570
Replication	"ch"			240	420
	"sh"			330	510
Experiment 2					
/tʃwæs/–/ʃwæs/	"ch"	142	232	221	376
	"sh"	242	333	322	476
Experiment 3					
/tʃkas/–/ʃkas/	"ch"	88	164	202	413
	"sh"	178	254	292	503
Experiment 4					
/twæs/–/swæs/	"t"	84	174	163	317
	"s"	220	310	299	453
Experiment 5					
/blos/–/plos/	"b" or "p"	83	135	174	337
/dlos/–/tlos/	"d" or "t"	106	158	197	360

quality (stylized, synthetic speech with few acoustic correlates and/or a noisy environment), then phonetic decisions may be delayed and a larger portion of the sound is used as the basis for rate normalization. If our /t/–/s/ series were of lower quality than our /tʃ/–/ʃ/ series, this post hoc explanation would work. Although there are no direct data available on this proposal, the results of Shinn, Blumstein, and Jongman (1985) and Miller and Wayland (1993) are suggestive. Shinn et al. showed that as a synthetic /b/–/w/ series became more natural sounding, the effect of the following vowel duration on the category boundary decreased and disappeared. Miller and Wayland showed that even high-quality stimuli produced the typical rate normalization result if presented against a background of other sound. Together, these results are consistent with the proposal that the scope of rate normalization effects for information following a target depends on the quality (and processing speed) of the acoustic-phonetic information specifying target identity.

Unfortunately, we have no easy way of testing this. Although we can compare the rating data from our listeners in the two experiments, it is unclear exactly what these results would indicate. These were two different groups of listeners, and the listeners in each group were likely to make their ratings in comparison with the range of sounds that they were being tested upon. That is, even if the listeners rated a stimulus end point as being a 1 or a 6 (the best rating possible) in relation to the other sounds he/she was hearing, this does not mean that these stimuli were as high in their phonetic quality as the stimuli with a similar rating in another experiment. Thus it is not clear that a comparison of rating data can truly answer this question. However, with these caveats in mind, we performed a 2 (experiment: /tʃ/–/ʃ/, /t/–/s/) × 2 (phonemes varying: vowel, approximant) × 3 (series: short, medium, long) × 2 (end point: short /t/, /tʃ/ vs. long /s/, /ʃ/) ANOVA. The only significant effect was that of experiment ($F = 7.207, p < .01$). However, the /t/–/s/ series in Experiment 4 actually received *better* overall ratings than did the /tʃ/–/ʃ/ series in Experiment 2, not worse. There were also two marginal interactions. One was between series (s, m, l) and phoneme (/w/, /æ/), which is not particularly relevant to the question of differential stimulus quality between experiments. The other was an interaction between the experiment and the phoneme varying ($F = 3.175, p < .08$), representing the fact that the difference in ratings between the /t/–/s/ series and the /tʃ/–/ʃ/ series was larger when the vowel was varying than when the /w/ was varying. In both cases, however, the /t/–/s/ series received the better ratings. Overall, then, these results do not provide any evidence that the /t/–/s/ series stimuli were more ambiguous than their matching /tʃ/–/ʃ/ series counterparts.

Finally, it is possible that phonotactic legality may be playing a role. Perhaps effects from later occurring phonemes can only occur in phonotactically legal series, even when they occur within the temporal window. Phonotactics alone cannot explain the results; the /tʃæs/–/ʃæs/

series from Experiment 1 was phonotactically legal, and it showed no effect of the duration of a distal phoneme. Furthermore, the effects from the distal vowel in this series (Experiment 4) only seemed to occur in the /t/ end of the series, even though both ends were legal in English. But perhaps there is an interaction between the temporal window and phonotactic legality, such that more distal sounds only have an effect if they meet both criteria (being legal and being temporally close). In order to examine this possibility, we decided to perform a direct test of the role of phonotactic legality.

EXPERIMENT 5

In this experiment, we created two series: one that ranged from /blos/ to /plos/ (the phonotactically legal series), and a second that ranged from /dlos/ to /tlos/ (the phonotactically illegal series). As in our other conditions, we varied the durations of both the liquid and the vowel portions of the syllables. We used the same /los/ portions for both series, so that the acoustic information and duration variation in the /l/ and /o/ portions were identical in the two series. Thus, the main difference separating the two series was one of phonotactic legality. If the two series should yield identical results, it would strongly suggest that phonotactic legality per se had little effect on rate normalization in this context. If the results should differ between the two series, it would suggest that phonotactics did play a role.

Both of these series are shorter than the ones used in the previous experiments. Thus, if there is a temporal window over which normalization can occur, we might expect to find an effect of duration variation of the distal vowel in these series. That is, the vowel in these series is closer to the initial phoneme, for both ends of the test series, than was the vowel in earlier series. This is shown in the bottom part of Table 8. If there is a relatively short temporal window for rate normalization, the vowel in these series might fall within it. The two series are very similar in their duration, so a temporal window account would predict similar results for both series.

Thus, these series allow us to examine the role of phonotactic legality directly. If the effect of the vowel in the /t/ end of the /twæs/–/swæs/ series was simply because of a temporal window, we would expect to see effects of the vowel in both the /dlos/–/tlos/ and the /blos/–/plos/ series. If our earlier results reflected effects of phonotactic legality, along with the temporal window, we would expect to see an influence of vowel duration only in the (legal) /blos/–/plos/ series.

Method

Subjects. The listeners were 77 students from an introductory psychology course at the State University of New York at Buffalo who participated in the experiment for class credit. All listeners were native speakers of English and had no reported history of a speech or hearing impairment. Thirty-nine students heard the /blos/–/plos/ series, of which 17 listeners heard the three sets of stimuli that varied in the liquid /l/ duration, and 22 heard the sets

with the variation in vowel /o/ duration. Thirty-eight listeners heard the /dlos/–/tlos/ series, of which 16 heard the sets with /l/ varying and 22 heard the sets with the /o/ varying. The data of 4 listeners were omitted from the analysis of the /blos/–/plos/ series for data loss (2 each from the liquid and vowel series). Also, 8 listeners (2 in the /blos/–/plos/ liquid series, 4 in the /blos/–/plos/ vowel series, 1 in the /dlos/–/tlos/ liquid series, and 1 in the /dlos/–/tlos/ vowel series) were dropped for a failure to categorize one or more end points accurately. This left a total of 13 listeners in the /blos/–/plos/ liquid condition, 16 in the /blos/–/plos/ vowel condition, 15 in the /dlos/–/tlos/ liquid condition, and 21 in the /dlos/–/tlos/ vowel condition.

Stimuli. A female native speaker of English (R.S.N.) recorded the syllables /blos/, /plos/, /dlos/, and /tlos/ in the context of running speech. We changed to a female voice because the higher fundamental frequency allowed for more fine-grained editing than did the lower fundamental of the male speaker. The stimuli were amplified, low-pass filtered at 9.5 kHz, digitized via a 12-bit, analog-to-digital converter at a 20-kHz sampling rate, and stored on computer disk. The syllables were excised from the carrier sentence, "Norton said _____ to me." The initial stop consonant segments were then separated from the remainder of the syllables. We considered the first 10 pulses as belonging to the /b/ and the first 12 as belonging to the /d/, basing this decision on examination of the formant transitions, particularly the first formant, within the syllable. A nine-member series ranging from /b/ to /p/ was created by replacing successively longer sections from the onset of the original /b/, up to the zero-crossing that marked the onset of a vocal pulse, with similar sections from the aperiodic /p/ onset. Similarly, a nine-member series ranging from /d/ to /t/ was created. Details of the waveform editing process for creating voicing series from natural speech can be found in Ganong (1980). This process of editing results in a natural-speech-based VOT continuum where the voicing in the initial /b/ or /d/ is gradually replaced by the aspiration of the initial /p/ or /t/. For both continua, the difference in VOT from stimulus to stimulus varies slightly since the voice pitch is not constant over the initial part of either /b/ or /d/. The size of the replaced sections averaged 7 msec for both the /b/–/p/ and /d/–/t/ series. The resulting series of VOTs ranged from approximately 24 to 61 msec in duration for the /b/–/p/ series, and 23 to 79 msec in duration for the /d/–/t/ series.

The remainder of the syllable, /los/, taken from the production of /blos/, was edited to create four new syllables: one with a short /l/, one with a short vowel, one with a long /l/, and one with a long vowel. We considered the first 11 pitch pulses as belonging to the /l/ for editing purposes, based on the movement of the formants. The 12th pulse seemed to belong equally to the /l/ and the /o/, so this boundary pulse was not considered part of either the liquid or the vowel. The remaining 38 pulses were all considered to belong to the /o/. However, the last 9 pulses of the vowel were not reduplicated in making the long vowel /los/, for fear of too greatly altering the amplitude envelope and formant transitions leading into the final /s/.

For the short /l/, every other pulse (the even-numbered pulses) was removed, for a total of 5 fewer pulses. For the long /l/, each of the pulses, except for the very first, was reduplicated, resulting in a stimulus with 10 more pulses. For the short /o/, every other pulse was deleted, for a total of 19 fewer pulses, and for the long /o/, each of the first 29 pulses was reduplicated. The short and long vowel durations both occurred with the intermediate liquid duration (38 msec), and the altered short and long /l/ stimuli had the intermediate vowel duration (142 msec). These five tokens of /los/ were then spliced to the ends of each of the nine members of the /b/–/p/ VOT continuum and the nine members of the /d/–/t/ continuum, resulting in 90 different syllables. Four sets of stimuli were created, one with the /l/ varying in the /blos/–/plos/ series (consisting of the short /l/ series, the intermediate series, and the long /l/ series), one with the /l/ varying in the /dlos/–/tlos/ series, one with the vowel varying in the /blos/–/plos/ series (consisting of the short

/o/ series, the intermediate series, and the long /o/ series), and one with the /o/ varying in the /dlos/–/tlos/ series. Each set had a total of 27 stimuli, and listeners heard only one of the four sets. The approximate durations for the liquid and vowel portions of these syllables are given in Table 9. These durations are within the range reported by Crystal and House (1988d) for liquids and vowels in American English.

Procedure. The procedure was nearly identical to that used in the previous experiments. All listeners heard a practice block containing two occurrences of each of the 27 syllables in their set in random order. This was followed by blocks of either 2 or 3 repetitions of each stimulus. All listeners heard at least 15 repetitions of each stimulus. Because of some hardware failures on individual blocks, the number of repetitions that each listener received was not identical. In the /blos/–/plos/ liquid series, 7 listeners heard seven blocks of three repetitions each, for a total of 21 repetitions, and the remaining 6 heard five blocks, for a total of 15. In the vowel series, 10 heard seven blocks (21 repetitions), and the remaining 6 heard five blocks (15 repetitions). Because of computer memory limitations, listeners heard 2 repetitions per block for the /dlos/–/tlos/ series. All listeners in this experiment heard eight blocks, or 16 repetitions of each stimulus.

Results and Discussion

For each listener, an average (mean) rating was computed for each stimulus in each series. The category boundary for each listener was then determined for each series, and a one-way ANOVA was run on these data for each of the four groups. As in previous experiments, an ANOVA was also performed on the percentages of voiced ("b" or "d") responses for each series. Table 10 shows the mean boundary locations and mean percentages (along with standard deviations) for the various conditions in this experiment.

We will consider the /b/–/p/ results first. As expected, the duration of the liquid had a significant effect on the initial /b/–/p/ distinction [$F(2,24) = 8.27, p < .005$]. Two of the subsequent paired comparisons were significant: short versus intermediate [$t(12) = 3.31, p < .05$], and short versus long [$t(12) = 3.30, p < .05$]. There was a marginal effect for the difference between the intermediate and long duration groups [$t(12) = 1.89, p < .10$]. With the percentage data, the results were similar. The overall effect was significant [$F(2,24) = 12.18, p < .0001$], as were all of the paired comparisons: short versus intermediate [$t(12) = 5.84, p < .001$], short versus long [$t(12) = 8.06, p < .001$], and intermediate versus long [$t(12) = 7.02, p < .001$].

The effect of vowel duration was also significant for the /b/–/p/ series [$F(2,30) = 10.20, p < .001$ for the cat-

Table 9
Liquid and Vowel Durations (in Milliseconds)
in the /blos/–/plos/ and /dlos/–/tlos/
Natural Speech Series of Experiment 5

Series	/l/ Varying		/o/ Varying	
Short	/l/	21	/l/	38
	/o/	142	/o/	74
Intermediate	/l/	38	/l/	38
	/o/	142	/o/	142
Long	/l/	73	/l/	38
	/o/	142	/o/	237

egory boundary data, and $F(2,30) = 18.66, p < .001$ for percentage data]. For the category boundary data, the subsequent paired comparisons showed significant effects of vowel duration for the short versus intermediate [$t(15) = 2.80, p < .01$] and short versus long [$t(15) = 1.83, p < .05$] pairs, but not for the intermediate versus long pair [$t(15) = 0.01, p > .10$]. For the percentage data, all the paired comparisons were significant: short versus intermediate [$t(15) = 3.58, p < .01$], short versus long [$t(15) = 4.82, p < .01$], and intermediate versus long [$t(15) = 4.89, p < .01$]. The mean rating functions, averaged across listeners, are shown in Figure 6. The data for the /l/ duration-varying group are on the left; the data for the /o/ duration-varying group are on the right.

For the /d/–/t/ series, the duration of the liquid had a significant effect on the initial /d/–/t/ distinction [$F(2,28) = 14.32, p < .0001$ for the category boundary data]. All three of the subsequent paired comparisons were significant: short versus intermediate [$t(14) = 2.38, p < .01$], intermediate versus long [$t(14) = 4.16, p < .001$], and short versus long [$t(14) = 4.52, p < .001$]. With the percentage "d" data, the results were similar. The overall effect was significant [$F(2,28) = 13.89, p < .0001$], as were all of the paired comparisons: short versus intermediate [$t(14) = 2.90, p < .01$], short versus long [$t(14) = 2.89, p < .01$], and intermediate versus long [$t(14) = 4.57, p < .001$].

The effect of vowel duration was also significant [$F(2,40) = 12.47, p < .001$ for the category boundary data, and $F(2,40) = 15.74, p < .001$ for the percent "d" data]. All of the paired comparisons were significant: short versus intermediate [$t(20) = 4.29, p < .001$], intermediate versus long [$t(20) = 2.09, p < .05$], and short versus long [$t(20) = 4.20, p < .001$] for the category boundaries; and short versus intermediate [$t(20) = 4.33, p < .001$], intermediate versus long [$t(20) = 2.78, p <$

.01], and short versus long [$t(20) = 4.88, p < .001$] for the percent "d."[6] The mean rating functions, averaged across listeners, are shown in the left panel of Figure 7 for the /l/ duration-varying group. The /o/ duration-varying group rating functions are shown on the right.

These results demonstrate that nonadjacent segments can have an effect on perception of an initial phoneme. Furthermore, they suggest that the primary determinant of whether a distal phoneme has an effect or not is whether it is within a critical temporal distance after the target. Segments that occur outside of (after) this critical distance do not appear to affect the rate normalization process. Furthermore, the rate normalization process does not appear to be influenced by whether the phonetic sequence is phonotactically legal or not. The only difference between our /blos/–/plos/ and /dlos/–/tlos/ series was in the initial consonant. This difference produced series that were either phonotactically legal for American English (/blos/–/plos/) or phonotactically illegal (/dlos/–/tlos/). However, robust effects of both the adjacent liquid duration and the remote vowel duration were found on both test series, indicating that no effect of phonotactics appears to have occurred. Thus, it would appear that the results of Experiments 2 and 3 were not influenced by the phonotactic status of the phoneme sequences that we used. Instead, we appear to have consistent evidence that only temporal proximity to the target is important for rate normalization.

GENERAL DISCUSSION

As an aid to discussing the results of all five experiments, Figure 8 shows the magnitude and direction of any change in categorization of the initial segment (/tʃ/–/ʃ/ for Experiments 1–3, /t/–/s/ in Experiment 4, and /b/–/p/ and /d/–/t/ in Experiment 5) in each experiment. Each pair of bars represents one of the six pairs of series, with the darker bar on the left representing the effect of variation in duration of the adjacent segment and the lighter bar on the right representing the effect of duration variation in the distal segment. Each bar shows the change in the overall percent classification between the short series and long series.

The presentation of the data in Figure 8 makes two points clear. First, the variation in duration of the adjacent phoneme, following the target, always produced a change in categorization of the initial segment (/tʃ/–/ʃ/, /t/–/s/, /b/–/p/, or /d/–/t/). The consistency of this effect is shown by the dark bars in Figure 8. Second, variation in the duration of the nonadjacent phoneme only produced a change in the /twæs/–/swæs/, /blos/–/plos/, and /dlos/–/tlos/ series. Analysis of the effect of the vowel in these three series showed that when the vowel was most temporally removed from the onset of the initial phoneme (/swæs/), no rate normalization effect was found (see also Table 8). This suggests that it is certainly possible to get rate normalization effects from nonadjacent segments, but only when they are within a limited temporal period after the onset of the target distinction.

Table 10
Experiment 5 Results

	Boundary Location (in msec)		Percentage "b" or "d"	
	M	SD	M	SD
/blos/–/plos/ Series				
/l/ liquid varying				
short series	34.9	4.05	46.62	6.35
intermediate series	37.4	1.70	48.97	13.94
long series	38.4	2.44	56.05	3.68
/o/ vowel varying				
short series	29.6	12.02	38.38	16.27
intermediate series	37.2	2.57	51.21	7.00
long series	39.6	2.41	56.91	6.55
/dlos/–/tlos/ Series				
/l/ liquid varying				
short series	38.3	4.81	47.25	7.32
intermediate series	40.0	4.97	49.34	7.65
long series	41.9	4.51	51.83	6.07
/o/ vowel varying				
short series	35.4	7.02	45.20	7.55
intermediate series	39.5	6.16	48.48	8.06
long series	42.9	10.57	53.27	11.08

Figure 6. Group rating functions for the /blos/–/plos/ series as a function of variation in glide duration (left) and vowel duration (right).

This overall pattern of results seems to be most readily explained by proposing that rate normalization based on information that follows the target occurs within a limited time window. Anything within that window can have an effect, but anything outside that window does not. Furthermore, it appears that the phonotactics of the utterance have little or no effect. In our series, rate normalization occurred in the same manner, whether the sequence was legal (e.g., /blos/) or illegal (/dlos/) in the language. Furthermore, across the limited range of phonetic contexts and test series used in these experiments, all adjacent phonemes (stop, approximants, vowel) produced an effect. To put this another way—we found no evidence that either phonetic identity or acoustic similarity of phonemes following the target influences rate normalization, in spite of our attempts to find such influences.

Our results are consistent with Miller and Dexter's (1988) findings and conclusions in two ways. First, there was no situation in which our listeners ignored later occurring rate information from the adjacent segment. Listeners consistently used rate information from the adjacent segments, without regard to segment identity (vowel, approximant, or stop) or phonotactics. To put this another way—our findings are consistent with Miller and Dexter's proposal that rate normalization is an obligatory component of speech processing and that listeners incorporate the rate information available at the time at which they make a phonetic decision. Second, the finding that neither phonotactics nor segment identity of the adjacent segment had an influence on rate normalization in these experiments is consistent with the idea that rate normalization occurs relatively early in perception.
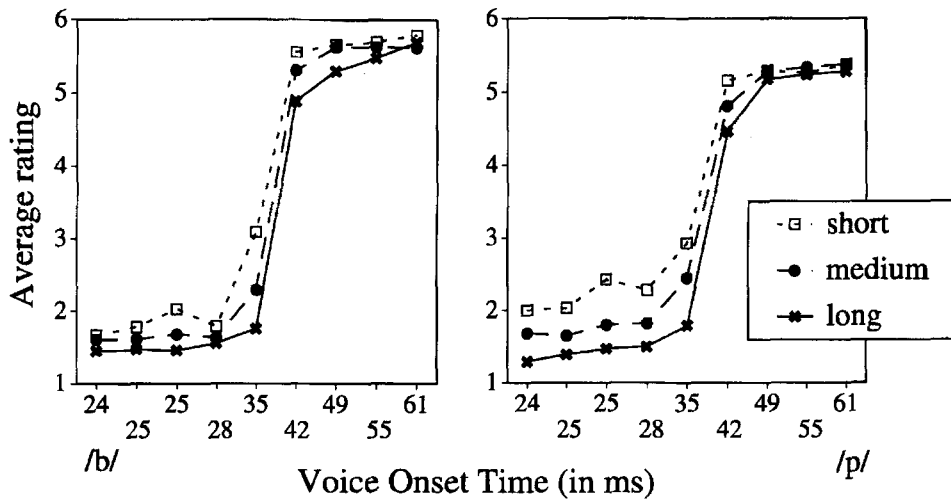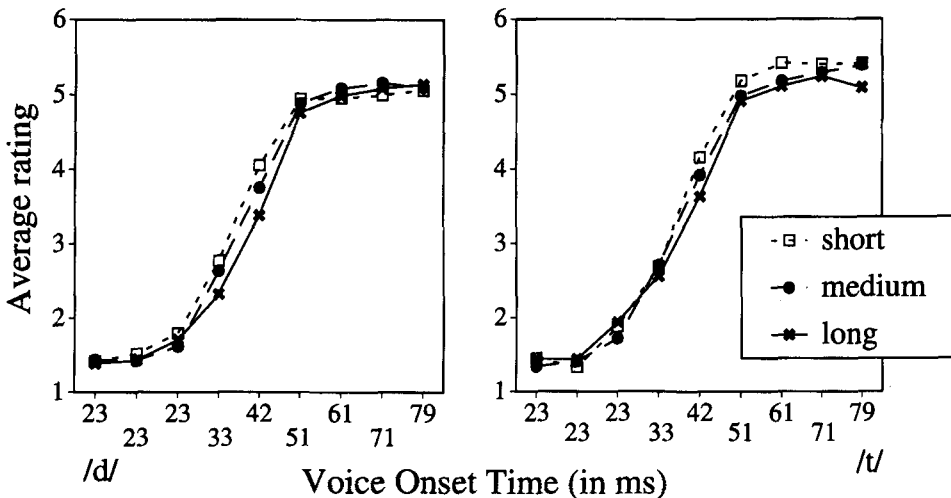

Figure 7. Group rating functions for the /dlos/–/tlos/ series as a function of variation in glide duration (left) and vowel duration (right).
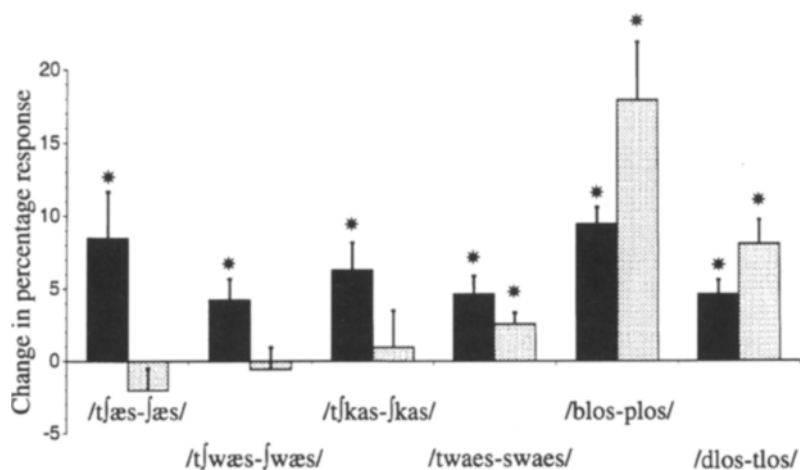
Figure 8. The magnitude of change in response to initial phoneme between short and long series in Experiments 1–5. In each pair of bars, the left, darker bar represents the effect of duration variation in the adjacent segment, and the right, lighter bar represents the effect of the more distal segment. *Significant effects. Error bars, standard error.

One explanation for these results is that there is a temporal window over which rate normalization initially occurs. This window normally includes one, or perhaps two, phonemes following the target segment. The duration of phonemes falling within this window will have an effect on perception of the target segment. That is, these durations will be the basis for rate normalization. Segment durations of phonemes falling outside of (after) this temporal window will not have such an effect. Furthermore, as described previously with respect to our results in Experiment 4 with the /twæs/–/swæs/ series, it is possible that the size of this temporal window is not fixed. Rather, it may vary with the time course of acoustic-phonetic processing. The time course of acoustic-phonetic processing, in turn, is influenced by the quality of the acoustic-phonetic information. With high-quality information, phonetic decisions are made relatively quickly and a relatively short temporal window would be present for rate normalization. Lower quality stimuli, or a noisy communication channel, may engender longer perceptual processing before a phonetic decision is reached. This yields a longer temporal window within which segments of the speech stream could influence rate normalization.

Although this proposal for a dynamic temporal window for rate normalization is post hoc, it is consistent with our data and with the effects of naturalness (Shinn et al., 1985) and of background noise (Miller & Wayland, 1993) on rate normalization. It also allows us to explain the remote rate normalization results reported by Miller and Liberman (1979) where the vowel in the syllable following the target produced rate normalization effects. In their study, duration variation in the second vowel in a /bada/–/wada/ series influenced perception of the initial phoneme as "b" or "w." The second vowel was relatively distant from the onset of the target when the first vowel was long (296 or 440 msec for short and long second vowels). When the first vowel was short, the time from target onset to the end of the second syllable was shorter (152 or 296 msec). Since Miller and Liberman found overall effects of both first vowel duration and second vowel duration, and no interaction, it would appear that information as much as 400 msec after target onset influenced perception. In our studies, no influence was found beyond about 300 msec from target onset. However, the Miller and Liberman stimuli were stylized, synthetic speech, whereas our stimuli were mostly edited, natural speech. If acoustic-phonetic processing for their stimuli took place over a slightly longer time course than that for our stimuli, there could have been a longer temporal window for rate normalization effects, and thus a consistent interpretation of all of these results emerges.

The data of Summerfield (1981) are also relevant here. In his studies, variation in the duration of a final /z/ fricative had no effect on the perception of an initial /b/–/p/ distinction in the syllables /biz/ and /piz/. The approximate duration from target onset to the end of the longest /z/ was 300 msec, or about the same as that to the end of the vowel in our /blos/–/plos/ stimuli. The fact that Summerfield found no effect of a nonadjacent consonant (replicated in Experiment 1 here), whereas we found an effect of a nonadjacent vowel when the temporal distances were roughly equal, is inconsistent with our proposal. Consequently, it is possible that additional factors are at work. One possibility is that the vowel between the target and the final /z/ exerted a strong enough effect so that it overwhelmed the influence of the /z/. A second possibility is that there is some role for a sonority principle, even though we did not observe it in the present experiments. A third possibility is that the listeners in Summerfield's study were making their phonetic decisions rapidly enough so that the final /z/ duration variation had no influence.

Finally, we should consider the results of Green et al. (1994) again. Their results showed that a substantial change in *F*0 stopped the rate normalization process. That is, only the vowel duration up to the *F*0 change in their CV syllables influenced perception of the initial target consonant. Like the results of Summerfield (1981), this suggests that some factor(s) besides temporal proximity contributes to rate normalization and that further research is needed.

In summary, the composite results of the experiments presented here suggest that rate normalization effects are dependent on a principle of adjacency, but not on principles of similarity or sonority. The phonotactics of a segment also do not appear to affect rate normalization. This adjacency principle is not limited to the adjacent phoneme per se, but rather to an adjacent temporal window, into which one or two phonemes typically fall. Only the segments that fall within that window are likely to have an effect on the rate normalization processes that operate in perception.

REFERENCES

AINSWORTH, W. A. (1977). Mechanisms of selective feature adaptation. *Perception & Psychophysics*, **21**, 365-370.

CRYSTAL, T. H., & HOUSE, A. S. (1982). Segmental duration in connected-speech signals: Preliminary results. *Journal of the Acoustical Society of America*, **72**, 705-716.

CRYSTAL, T. H., & HOUSE, A. S. (1988a). The duration of American-English stop consonants: An overview. *Journal of Phonetics*, **16**, 285-294.

CRYSTAL, T. H., & HOUSE, A. S. (1988b). The duration of American-English vowels: An overview. *Journal of Phonetics*, **16**, 263-284.

CRYSTAL, T. H., & HOUSE, A. S. (1988c). A note on the durations of fricatives in American English. *Journal of the Acoustical Society of America*, **84**, 1932-1935.

CRYSTAL, T. H., & HOUSE, A. S. (1988d). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America*, **83**, 1553-1573.

CRYSTAL, T. H., & HOUSE, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America*, **88**, 101-112.

CUTTING, J. E., & ROSNER, B. S. (1974). Categories and boundaries in speech and music. *Perception & Psychophysics*, **16**, 564-570.

DAVIDSEN-NIELSEN, N. (1974). Syllabification in English words with medial sp, st, sk. *Journal of Phonetics*, **2**, 15-45.

DIEHL, R. L., & WALSH, M. A. (1989). An auditory basis for the stimulus-length effect in the perception of stops and glides. *Journal of the Acoustical Society of America*, **85**, 2154-2164.

FOWLER, C. A. (1983). Converging sources of information for spoken and perceived rhythms of speech. *Journal of Experimental Psychology: General*, **112**, 386-412.

FOWLER, C. A. (1987). Perceivers as realists, talkers too: Commentary on papers by Strange, Diehl et al., and Rakerd and Verbrugge. *Journal of Memory & Language*, **26**, 574-587.

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

GAY, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America*, **63**, 223-230.

GERSTMAN, L. J. (1957). *Cues for distinguishing among fricatives, affricates, and stop consonants*. Unpublished doctoral dissertation, New York University.

GREEN, K. P., STEVENS, E. B., & KUHL, P. K. (1994). Talker continuity and the use of rate information during phonetic perception. *Perception & Psychophysics*, **55**, 249-260.

KIDD, G. R. (1989). Articulatory-rate context effects in phoneme iden-

tification. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 736-748.

KLATT, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech & Hearing Research*, **18**, 686-706.

KLATT, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971-995.

KLUENDER, K. R., & WALSH, M. A. (1988). Effect of vowel duration on the perception of syllable-initial /sh/ and /ch/. *Journal of the Acoustical Society of America*, **84**, S159.

KLUENDER, K. R., & WALSH, M. A. (1992). Amplitude rise time and the perception of the voiceless affricate/fricative distinction. *Perception & Psychophysics*, **51**, 328-333.

LIBERMAN, A. M., DELATTRE, P. C., GERSTMAN, L. J., & COOPER, F. S. (1956). Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology*, **52**, 127-137.

LOTTO, A. J., KLUENDER, K. R., & HOLT, L. L. (1994, November). *Spectral discontinuities and the vowel length effect*. Paper presented at the 128th meeting of the Acoustical Society of America, Houston.

MILLER, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39-74). Hillsdale, NJ: Erlbaum.

MILLER, J. L., & BAER, T. (1983). Some effects of speaking rate on the production of /b/ and /w/. *Journal of the Acoustical Society of America*, **73**, 1751-1755.

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., GROSJEAN, F., & LOMANTO, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica*, **41**, 215-225.

MILLER, J. L., & LIBERMAN, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, **25**, 457-465.

MILLER, J. L., & WAYLAND, S. C. (1993). Limits in the limitations of context-conditioned effects in the perception of [b] and [w]. *Perception & Psychophysics*, **54**, 205-210.

PISONI, D. B., CARRELL, T. D., & GANS, S. J. (1983). Perception of the duration of rapid spectrum changes in speech and nonspeech signals. *Perception & Psychophysics*, **34**, 314-322.

REPP, B. H., LIBERMAN, A. M., ECCARDT, T., & PESETSKY, D. (1978). Perceptual integration of acoustic cues for stop, fricative and affricate manner. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 621-637.

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

SAMUEL, A. (1988). Central and peripheral representation of whispered and voiced speech. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 379-388.

SAWUSCH, J. R. (1976). Selective adaptation effects on end-point stimuli in a speech series. *Perception & Psychophysics*, **20**, 61-65.

SHINN, P. C., BLUMSTEIN, S. E., & JONGMAN, A. (1985). Limitations of context conditioned effects in the perception of [b] and [w]. *Perception & Psychophysics*, **38**, 397-407.

SUMMERFIELD, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1074-1095.

NOTES

1. This shows that duration, per se, is not the only factor involved in rate normalization.

2. We considered an individual as being unable to classify the end points if he/she did not reach 80% correct at both ends of at least two of the three series with which he/she was tested. Across all five experiments, 31 of 232 listeners (13.4%) failed this criterion. The failure rate was higher for Experiments 1–3 (21 of 114, or 18.4%) than it was for Experiments 4 and 5 (10 of 118, or 8.5%). Since Experiments 1–3 all used /tʃ/–/ʃ/ test series, this may reflect a general difficulty that listeners had in classifying these stimuli. The reason for this failure rate is un-

clear. However, the general consistency of the results within and across experiments indicates to us that this failure rate should not substantially influence the conclusions we have drawn.

3. The percentage data were computed by collapsing across the ratings. Thus, ratings of 1, 2, and 3 were treated as "ch" responses, and ratings of 4, 5, and 6 were treated as "sh" responses.

4. For the planned comparisons, two-tailed t tests were used unless otherwise noted.

5. The /k/ in /ʃkas/ is only voiceless phonemically. The VOT of 31 msec is acoustically similar to that of a syllable initial /g/. This voice onset time is appropriate for a /k/ following a fricative. Davidsen-Nielsen (1974) showed that typical VOT values for /k/ following /s/ fricatives range from 20 to 50 msec, with an average VOT of 30 msec. Klatt (1975) found a VOT range for /k/ in /sk/ clusters of 25–39 msec for his 3 talkers, with an average of 30 msec. Thus, our talker's VOT for /k/ following /ʃ/ was very similar to the average found for /k/ following /s/ in prior research.

6. The first time we ran the /dlos/–/tlos/ series with the vowel varying, we did not find any significant effects [$F(2,28) = 1.14, p > .10$ for category boundary data, and $F(2,28) = 2.19, p > .10$ for percentage data]. However, there did appear to be a trend in the correct direction. Therefore, we ran the experiment a second time, and this time we did find significant effects; these are the results reported in the main text. When we combined the data from both experiments, the results were significant [$F(2,70) = 10.70, p < .0001$ by boundaries, and $F(2,70) = 15.09$, $p < .0001$ by percents]. The follow-up tests were also all significant [for categories, short vs. intermediate, $t(35) = 3.00, p < .005$; intermediate vs. long, $t(35) = 2.53, p < .01$; and short vs. long, $t(35) = 3.74, p < .0005$; for percent "d," short vs. intermediate, $t(35) = 4.42, p < .0005$; intermediate vs. long, $t(35) = 2.80, p < .005$; and short vs. long, $t(35) = 4.42, p < .0005$]. This suggests that the nonsignificant results from our first attempt were due to error variance, and that the significant results reported above are in fact the more reliable results.

APPENDIX

The formant frequencies, bandwidths, and amplitudes for the first through fifth formants, over the entire /tʃæs/ syllable, are shown below. The fundamental frequency ($F0$), amplitude of voicing (AV), and amplitude of frication (AF) are also shown. All other parameters of the synthetic stimuli in Experiment 1 were held constant. The amplitude of the sixth formant (A6) was set to zero. Other values were set to their defaults (see Klatt, 1980).

Parameter Listings for /tʃæs/

Milliseconds	F1	B1	A1	F2	B2	A2	F3	B3	A3	F4	B4	A4	F5	B5	A5	F0	AV	AF
0	377	240	11	1644	430	27	2349	161	44	3416	163	54	4226	150	36	105	0	40
5	377	245	11	1644	175	27	2351	164	44	3413	163	54	4223	148	36	105	0	49
10	377	374	11	1644	175	27	2352	167	44	3411	163	54	4221	146	36	105	0	57
15	377	144	11	1644	496	27	2360	186	44	3397	163	54	4207	134	36	105	0	58
20	377	144	11	1644	496	27	2360	186	44	3397	163	54	4207	134	36	105	0	58
25	377	256	13	1644	350	27	2370	430	44	3377	200	54	4187	118	36	104	0	58
30	377	180	20	1644	350	29	2374	145	43	3370	201	53	4180	111	34	108	0	58
35	377	163	23	1644	350	29	2375	222	43	3368	202	53	4178	109	34	110	0	58
40	377	213	25	1644	350	29	2376	136	42	3365	203	53	4175	107	34	113	0	58
45	377	310	27	1644	350	30	2378	242	42	3363	203	52	4173	105	33	115	0	58
50	377	375	30	1644	350	30	2379	95	42	3361	204	52	4171	103	33	117	0	53
55	377	245	32	1644	245	31	2401	51	41	3358	205	48	4168	101	32	119	0	48
60	406	167	34	1644	212	31	2356	68	41	3356	206	44	4166	99	32	122	0	43
65	434	152	37	1688	133	32	2386	127	41	3353	207	40	4163	97	32	122	32	0
70	464	99	39	1719	126	32	2450	80	41	3351	208	36	4161	95	31	123	37	0
75	500	72	41	1742	105	35	2462	73	40	3381	208	32	4191	93	31	123	42	0
80	522	58	44	1778	99	35	2462	77	40	3397	209	28	4207	91	30	124	47	0
85	528	45	47	1756	101	35	2483	64	41	3389	210	24	4199	89	30	124	51	0
90	528	41	49	1714	97	35	2496	69	40	3407	211	20	4217	87	29	124	55	0
95	530	43	49	1725	91	35	2518	73	39	3420	212	18	4230	85	32	124	55	0
100	541	47	49	1726	76	37	2516	71	40	3401	213	15	4211	83	35	124	55	0
105	549	56	46	1707	76	37	2506	77	39	3394	213	15	4204	81	33	123	56	0
110	561	55	46	1701	69	38	2506	76	39	3389	214	15	4199	80	33	123	56	0
115	567	50	46	1675	76	36	2500	76	39	3393	215	15	4203	93	32	122	56	0
120	569	54	45	1662	75	35	2518	68	37	3401	216	15	4211	99	28	122	56	0
125	585	62	44	1639	61	37	2489	98	34	3458	217	15	4268	130	23	121	57	0
130	603	64	42	1637	65	37	2471	105	35	3428	218	15	4238	110	23	120	57	0
135	606	53	44	1640	70	37	2487	89	37	3411	218	15	4221	124	25	119	57	0
140	623	61	43	1619	80	37	2475	70	41	3432	219	15	4242	91	29	118	57	0
145	626	76	40	1625	82	35	2466	61	40	3432	220	15	4242	85	27	117	57	0
150	635	100	37	1624	80	34	2470	63	40	3430	221	15	4240	90	27	116	58	0
155	658	110	35	1592	96	33	2473	78	38	3456	222	15	4266	97	26	115	58	0
160	639	104	36	1583	93	33	2481	65	39	3457	223	15	4267	86	27	114	58	0
165	663	101	36	1569	109	33	2498	68	39	3479	223	15	4289	110	27	113	58	0
170	691	139	34	1589	97	33	2488	69	38	3471	224	15	4281	113	25	112	58	0
175	704	109	36	1587	74	36	2478	84	37	3470	225	15	4280	84	28	111	59	0
180	700	164	34	1595	65	37	2479	90	35	3461	226	15	4271	99	24	110	59	0
185	686	110	36	1577	57	38	2484	66	39	3452	227	15	4262	93	24	109	59	0
190	670	119	36	1562	65	37	2495	73	38	3454	228	15	4264	85	27	109	59	0
195	683	101	36	1563	54	37	2515	70	37	3443	228	15	4253	82	26	108	60	0
200	672	120	35	1543	53	38	2544	76	36	3454	229	15	4264	119	23	108	60	0

APPENDIX (Continued)

Milliseconds	F1	B1	A1	F2	B2	A2	F3	B3	A3	F4	B4	A4	F5	B5	A5	F0	AV	AF
205	683	90	36	1532	56	38	2526	73	35	3455	230	15	4265	99	24	107	60	0
210	667	136	34	1523	55	37	2524	98	33	3440	231	15	4250	114	21	107	60	0
215	686	115	34	1524	45	37	2539	79	33	3443	232	15	4253	95	21	107	60	0
220	703	228	33	1517	70	34	2534	94	34	3449	233	15	4259	76	24	107	59	0
225	707	132	33	1511	61	34	2554	82	33	3463	233	15	4273	74	23	106	59	0
230	690	138	34	1488	75	33	2583	89	32	3502	234	15	4312	89	23	106	59	0
235	697	123	34	1484	57	34	2591	78	32	3497	235	15	4307	90	22	106	59	0
240	685	146	33	1497	57	34	2585	99	30	3508	236	15	4318	107	20	106	59	0
245	703	133	34	1511	51	35	2592	75	30	3529	237	15	4339	104	20	106	59	0
250	687	110	34	1524	63	33	2591	67	29	3569	238	15	4379	95	18	106	59	0
255	692	91	34	1537	42	33	2595	52	27	3580	238	15	4390	70	19	106	59	0
260	703	75	32	1550	62	25	2609	111	14	3592	239	15	4402	57	20	106	58	0
265	717	46	33	1564	78	19	2619	65	16	3588	240	15	4398	60	18	105	55	0
270	716	58	27	1577	101	13	2627	69	12	3542	241	15	4352	120	9	105	53	0
275	704	90	19	1590	66	18	2636	82	11	3618	242	15	4428	132	5	105	50	0
280	681	61	22	1603	93	13	2644	132	8	3563	243	15	4373	214	12	105	48	0
285	678	44	24	1617	58	13	2652	131	12	3607	243	15	4417	214	18	105	45	0
290	657	53	22	1630	86	16	2661	130	17	3670	244	15	4480	149	25	105	43	0
295	672	70	22	1632	124	19	2669	254	21	3451	245	15	4261	211	32	105	40	0
300	670	327	21	1633	250	23	2670	288	25	3511	246	37	4321	106	39	105	0	41
305	667	407	21	1635	167	26	2670	291	30	3511	246	37	4321	295	45	105	0	43
310	665	407	21	1636	341	29	2671	198	34	3511	246	37	4321	99	52	105	0	45
315	662	331	20	1638	341	29	2672	177	34	3511	246	37	4321	75	52	105	0	47
320	660	198	20	1639	351	29	2673	99	34	3510	246	37	4320	436	52	105	0	49
325	658	227	19	1641	245	29	2673	209	34	3510	121	37	4320	283	52	105	0	51
330	655	500	19	1642	190	29	2674	304	34	3510	62	37	4320	250	52	105	0	53
335	653	281	19	1644	331	29	2675	109	34	3510	55	37	4320	213	52	105	0	55
340	650	253	18	1646	258	29	2675	165	34	3509	148	37	4319	240	52	105	0	55
345	648	253	18	1647	260	29	2676	165	34	3509	261	37	4319	434	52	105	0	55
350	645	295	18	1649	466	29	2677	165	34	3509	213	37	4319	267	52	105	0	55
355	643	235	17	1650	333	29	2677	334	34	3509	241	37	4319	344	52	105	0	55
360	641	191	17	1652	333	29	2678	334	34	3508	217	37	4318	186	52	105	0	55
365	629	143	17	1660	421	29	2682	452	34	3507	202	37	4317	245	52	105	0	55
370	626	279	16	1661	421	29	2682	300	34	3507	270	37	4317	285	52	105	0	55
375	624	358	16	1663	421	29	2683	300	34	3507	270	37	4317	212	52	105	0	55
380	621	358	15	1664	421	29	2684	282	34	3506	388	37	4316	274	52	105	0	55
385	619	225	15	1666	368	29	2685	208	34	3506	388	37	4316	309	52	105	0	55
390	619	225	15	1666	368	29	2685	208	34	3506	388	37	4316	309	52	105	0	55
395	617	216	15	1667	226	29	2685	202	34	3506	351	37	4316	207	52	105	0	55
400	617	216	15	1667	226	29	2685	202	34	3506	351	37	4316	207	52	105	0	52
405	614	322	14	1669	336	29	2686	256	34	3506	419	37	4316	250	52	105	0	48
410	614	322	14	1669	336	29	2686	256	34	3506	419	37	4316	250	52	105	0	45
415	612	274	14	1670	388	29	2687	256	34	3505	448	37	4315	441	52	105	0	41
420	612	274	14	1670	388	29	2687	256	34	3505	448	37	4315	441	52	105	0	38
425	609	194	14	1672	309	29	2687	239	34	3505	247	37	4315	161	52	105	0	34
430	609	194	14	1672	309	29	2687	239	34	3505	247	37	4315	161	52	105	0	31
435	607	175	13	1674	342	29	2688	172	34	3505	247	37	4315	279	52	105	0	27