

P-CENTERS ARE UNAFFECTED BY PHONETIC CATEGORIZATION*

André Maurice Cooper,† D. H. Whalen, and Carol Ann Fowler‡

Abstract. The perceived onset (P-center) of a word typically does not correspond to its acoustic onset (Marcus, 1981; Morton, Marcus & Frankish, 1976). Some researchers have suggested that the P-center of a word is solely a product of the acoustic characteristics of the word, while others have suggested that a word's P-center is determined by its phonetic characteristics. The present series of experiments pits a continuously varying acoustic parameter against a categorical phonetic percept in order to determine whether P-center location is sensitive to the phonetic identity of the prevocalic segments of a syllable. With a /ʃa/-/ča/-/ta/ continuum and three different /sa/-/sta/ continua, we find that phonetic judgments are categorical but P-center judgments are continuous. The results demonstrate that P-center location is not determined by the phonetic identity of syllable initial consonants. Nor, however, is it determined by the rise time or the amplitude envelope of the signal as Howell (1984) has suggested. Instead, as Morton et al. and Marcus recognized, a combination of at least two different parts of the signal is at work, namely, the duration of the prevocalic consonant or consonants and, to a lesser extent, the duration of the syllable rhyme. Whereas the relevant dimension of each component of the syllable is duration, acoustically defined, the partitioning of the syllable is phonetically motivated. Thus, both the phonetic structure of a syllable and the particular acoustic realizations of its structure affect the location of the P-center.

When listeners are presented with sequences of consonant-vowel syllables differing in the number or the nature of consonants and with equal intervals between their acoustic onsets, they judge the rhythm of the sequences to be irregular. Furthermore, when given the opportunity to adjust the relative timing of two syllables until they are perceptually isochronous, listeners introduce systematic deviations from acoustic onset isochrony (Morton et al., 1976). These deviations can not be explained by reference to any obvious acoustic events such as the peak intensity of an utterance or the acoustically marked onset of the stressed vowel (see Allen, 1972; Morton et al., 1976; Rapp, 1971).

These findings indicate that the event that listeners attend to when judging relative timing is opaque to conventional measurement techniques.

*Perception & Psychophysics, 1986, 39, 187-196.

†Also Department of Linguistics, Yale University.

‡Also Department of Psychology, Dartmouth University.

Acknowledgment. This research was supported by NIH Grant 16591, NIH Grant HD 01994, and NSF Grant BNS 8111470 to Haskins Laboratories.

Therefore, Morton et al. (1976, p. 405) do not attempt to locate the event absolutely, but propose that listeners base their rhythmicity judgments on a word's "psychological moment of occurrence" or its "P-center" (perceptual center). According to Marcus (1981), although the P-center of a word cannot be determined absolutely, it can be located relative to the timing of other speech or nonspeech events. Thus, by hypothesis, in order for two syllables presented in continuous alternation to be perceived as isochronous, the components of the sequence must have their P-centers at equal intervals.

Syllables whose initial consonants differ in manner generally have different P-center locations (Fowler & Tassinary, 1981). These consonants, in turn, have different acoustic characteristics, especially in the duration of the signal before the first vocalic pitch period (loosely, "consonant duration"). Marcus (1981) found that the location of the P-center is highly correlated with initial consonant duration. Specifically, the shorter the duration of the consonant, the earlier the P-center with respect to the acoustic onset of the syllable (also see Rapp, 1971; Fowler & Tassinary, 1981). Marcus also found that changes in the duration of segments following the initial consonants are associated with a smaller yet significant change in the location of the P-center. These two relationships are expressed by the equation:

$$P = .65x + .25y + k,$$

where x is the measured duration of a syllable onset (that is, the part of the signal preceding the first oral pitch pulse), y is the duration of the syllable rhyme (the vocalic segment and final consonants) and k is an arbitrary constant reflecting the fact that the equation predicts the relative, rather than the absolute, location of the P-center. Although the equation accounts for about 90% of the variability of P-center locations in the set of digits one to nine, it does not explain this variability.

In fact, one issue that the equation leaves in question is whether P-center shifts are explained by the phonetic or by the acoustic properties of a word. According to Marcus's equation, P-centers have a phonetic basis to the extent that the equation predicts that syllable-initial consonants have a markedly greater affect on P-center location than do vowels (whether syllable initial or nonsyllable initial) or final consonants. However, Marcus also shows that durational changes that do not affect the phonetic identity of the segments in a word do affect P-center location. Accordingly, the P-center is not solely a product of the phonetic identity of a segment.

Marcus attempted to test the effect of phonetic identity on P-center location by pairing the members of a phonetic continuum with several reference stimuli and adjusting the relative timing of the stimulus pairs to isochrony. The continuum was created by deleting successive portions of the /s/-noise (in 30-ms decrements) from a naturally-produced token of the word /sevən/. The initial consonant of the continuum stimuli spanned three phonetic categories, viz., /s/, /ts/ and /d/ (presumably as judged by Marcus, 1976, himself). Marcus found that abrupt shifts in phone categorization across the phonetic continuum were accompanied by a continuous change in P-center location across the continuum. Marcus's continuum, however, was not sufficiently well constructed to address the issue adequately. First, no identification or discrimination tests were performed to confirm how the continuum was perceived. Second, the steps of Marcus's continuum were so large that the

initial consonants spanned three phonetic categories (including one, /ts/, that is not phonotactically possible in English) within four steps of the five-step continuum.

In the present study, we used a series of phonetic continua to investigate the extent to which the phonetic realization of a syllable-initial consonant is relevant to P-center location. Our experiments improved upon and extended Marcus's experiment in several ways. First, we obtained identification and discrimination data to confirm that the continuum stimuli were categorically perceived. Second, the acoustic differences between neighboring stimuli on the continuum were sufficiently small to enable us to address the question of the relationship between phonetic consonantal categories and the location of the P-center. Third, all of the phonetic categories under investigation, /ʒ/-/ʒ/-/t/ and /s/-/st/, occur syllable-initially in English, the native language of the listeners. The phonetic categories /s/ and /st/ had already been shown to have P-center differences on the order of 26 ms (Fowler & Tassinary, 1981). Finally, we manipulated both prevocalic and postvocalic durations, allowing us to test Marcus's equation directly. Our results also allowed us to address Howell's (1984) recent suggestion that the amplitude envelope of a syllable significantly affects the location of its P-center.

We compared the phonetic categorization of syllable initial consonants and the relative location of P-centers to determine their effects upon each other. The test stimuli consisted of a /ʒa/-/ʒa/-/ta/ continuum and three /sa/-/sta/ continua. The construction of our first continuum was guided by the fact that when a sufficient amount of friction is deleted from /ʒa/, listeners hear /ʒa/ rather than /ʒa/; as additional friction is deleted, listeners eventually hear /ta/. The construction of the remaining continua was based on the fact that when a sufficiently long silent gap is introduced between the friction and the vocalic segment of the /sa/ syllable, listeners hear /sta/ rather than /sa/.

Repp (1984) identified four criteria that delimit categorical perception: there must be (1) an abrupt shift in labeling probabilities somewhere along the continuum, (2) a peak in the discrimination function at the category boundary, (3) chance or near-chance level discrimination of stimuli within categories, and (4) perfect predictability of the discrimination function from the identification function. Strict categorical perception, as described above, is rarely, if ever, reported in the literature; instead, the actual data approximate the ideal more or less well. Repp (1984) emphasizes that provided that the other criteria are not severely violated, a peak in the discrimination function at the category boundary is the crucial defining characteristic of categorical perception.

Three relationships between the categorical nature of the continua and the P-center function are possible. First, the relative location of the P-center could be sensitive only to the phonetic properties of the syllables. In this case, listeners' P-center judgments would be predictable from their identification functions. For syllables with categorically-perceived consonants, this would imply negligible within-category P-center shifts, but noticeable shifts between categories. Second, the P-center could be sensitive to the durations of a syllable's acoustic segments. In this case, listeners' P-center judgments would vary monotonically as a function of duration (see Marcus, 1981). The third possibility is that P-center judgments would be

influenced by both phonetic and acoustic properties of the stimuli, resulting in both an abrupt shift in the P-center at the category boundary and systematic variation within categories.

General Procedures, Experiments 1-4

Each experiment consisted of three tasks. The first two tasks, a forced-choice identification test and an AXB discrimination test, were used to determine the extent to which each continuum was categorically perceived. For the identification test, multiple repetitions of the stimuli were randomized and presented to listeners. For the AXB test, multiple repetitions of all pairs of syllables in a continuum differing by one step (Experiment 1) or by two steps (Experiments 2 - 4) were randomized and presented to listeners.

The final task, the alignment test, was designed to measure the relative P-center location of the test stimuli. For the alignment test, each of the test stimuli was paired with a reference syllable /ba/ for presentation. (The reference syllable was 329 ms in duration). The syllable pairs were played to listeners in a continuous sequence under computer control. The temporal position of the second syllable relative to the first was adjustable within a window of fixed duration. Initially, on each trial, there was a 50-ms gap between offset of the fixed syllable and the onset of the movable stimulus. The listener's task was to adjust the timing of the sequence until it was perceived as isochronous. When the listener was satisfied with the adjustment, the computer reported the interval between the acoustic onsets of the stimuli. Two systems, with minor differences, were used. On one system, implemented with a New England Digital computer at Dartmouth College, listeners adjusted the second syllable in steps of 15 ms, 5 ms or 1 ms in either direction relative to the fixed syllable by pressing designated keys on a computer terminal keyboard. On the other system, implemented with a DEC GT40 computer at Haskins Laboratories, the alignments were made by turning a knob. The analog output of the knob was digitized to indicate adjustments in 12.8-ms increments. Since the experimental results obtained from the two systems were similar, they were combined.

Experiment 1

The purpose of the first experiment was to investigate how the location of P-centers might vary across a categorically perceived /ʒa/-/ʒa/-/ta/ continuum.

Method

Stimuli. Using a waveform editor, a 10-step /ʒa/-/ʒa/-/ta/ continuum was created by deleting 15-ms increments of frication from the acoustic onset of a digitized naturally spoken /ʒa/ syllable. The fricative segment of the original /ʒa/ was 189 ms in duration; the vocalic segment was 266 ms in duration. Syllable duration covaried with the duration of /ʒ/ in the continuum. To minimize abrupt onsets, an amplitude ramp, linear with sound pressure, was applied to the onset of each stimulus. The offset of the ramp was fixed at 150 ms into the frication of the original /ʒa/, while the total duration (and steepness) of the ramp varied with the duration of the frication. Thus, for one extreme of the continuum, the ramp was applied to the initial 150 ms of the original stimulus. For the other extreme, the first 135 ms of noise was deleted from the frication and a linear taper was applied

to the initial 15 ms of the frication (see Figure 1). Although the onset ramps in our continua do become steeper as the continuum stimuli get shorter, this seemed to us an improvement over the procedure of Marcus (1981), who made no attempt to avoid abrupt onsets. In his continuum, increasing amounts of fricative energy were simply removed from the beginning of the word "seven," resulting in abrupt onsets, and hence the perception of /ts/ very early in the continuum.

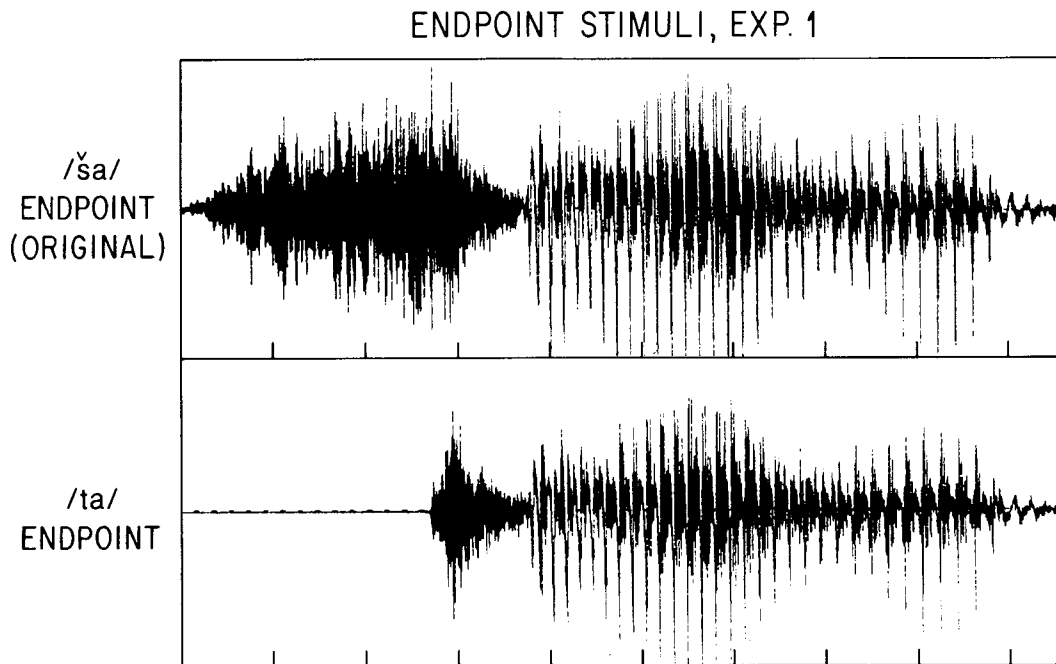


Figure 1. Continuum endpoint stimuli for Experiment 1: unmodified stimulus (upper panel), extreme stimulus manipulation (lower panel).

The identification test consisted of a randomized sequence of 10 repetitions of each member of the continuum (10 x 10 = 100 trials). The categories allowed in the identification test were /t/, /ʃ/ and /ʒ/. A one-step AXB discrimination test consisted of a randomized sequence of five repetitions of the four versions of each of the nine pairings of the stimuli (9 x 4 x 5 = 180 trials). For the alignment test, twenty judgments were obtained for each member of the continuum (20 x 10 = 200 trials).

Subjects. Three subjects participated. Two were naive as to the purposes of the experiment and the third was one of the authors (CAF).

Results and Discussion

Figure 2 shows the mean identification and discrimination functions for the three subjects. The ordinate represents the percent identification and the percent of correct discrimination. The abscissa represents the members of the continuum.

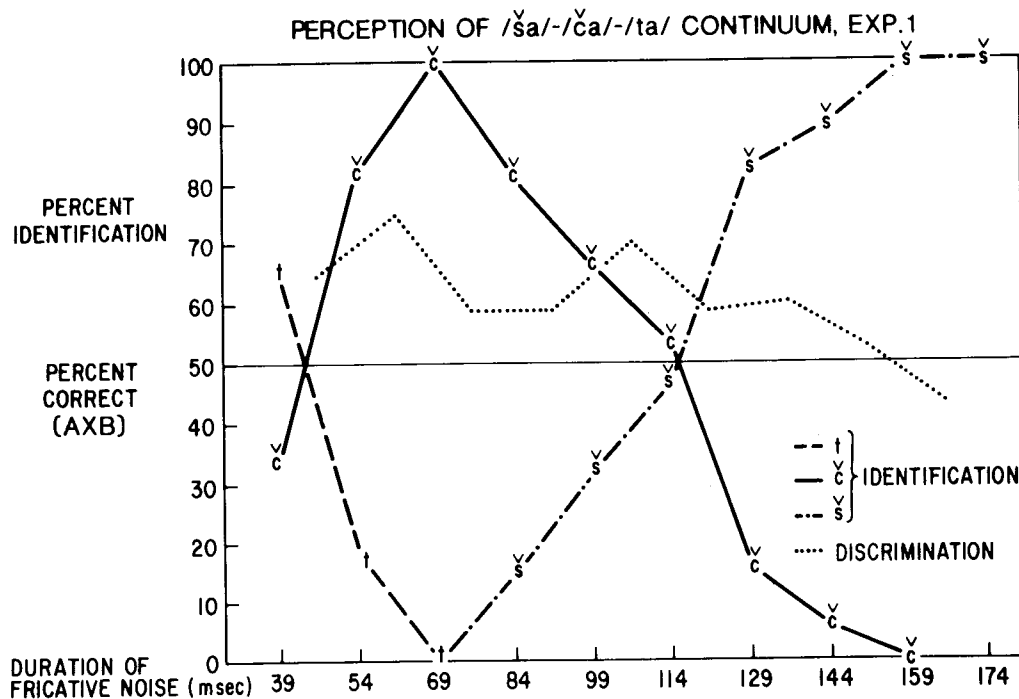


Figure 2. Mean identification and discrimination functions (pooled across three subjects) for /ʃa/-/ʧa/-/ta/ continuum, for Experiment 1.

The identification function shows two abrupt shifts in phoneme identification. The mean category boundaries (the 50% crossover points of the identification function) occur at 115 ms of frication for /ʃa/-/ʧa/ for all subjects and at 44 ms of frication for /ʧa/-/ta/ for the two subjects who reported /ta/'s. The discrimination function shows peaks in discrimination near the mean category boundaries, indicating that listeners discriminate better between stimuli that straddle a phoneme boundary than between stimuli that fall within the same phoneme category. Our data show two departures from strict categorical perception that are frequently reported in the literature. First, our discrimination peak is slightly offset from the category boundary, and second, within-category discrimination is above chance level (see Best, Morrongiello, & Robson, 1981; Healy & Repp, 1982; Liberman, Harris, Eimas, Lisker, & Bastian, 1961; Liberman, Harris, Hoffman, & Griffith, 1957; Liberman, Harris, Kinney, & Lane, 1961). Nevertheless, the pattern of our data is similar to the patterns obtained in earlier studies in which researchers concluded that perception was categorical.

The mean results of the three subjects' performance on the alignment test are shown in Figure 3. The ordinate represents the displacement from acoustic isochrony of the test stimuli relative to the reference syllable in ms (i.e., the interval from the acoustic onset of the test stimuli to the acoustic onset of the reference syllable minus one half the window size; thus, this measure would be zero for stimuli aligned at their acoustic onsets). The abscissa represents the duration of the fricative noise in ms.

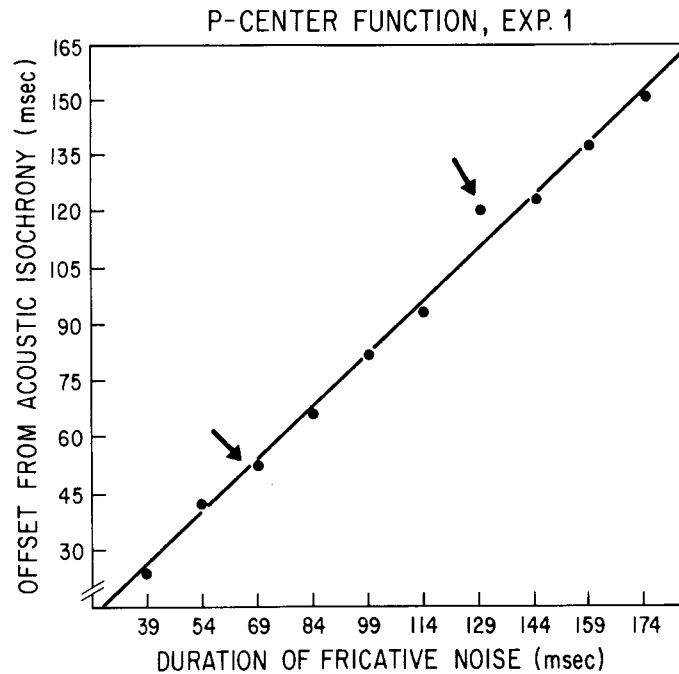


Figure 3. Mean P-center alignment function for Experiment 1, pooled across three subjects. The solid line represents the regression line and the arrows indicate the category boundaries.

In this experiment, P-center location, as shown by the regression line, moves linearly toward stimulus offset as the duration of the noise increases across the continuum. The slope of the regression line is .95, with $r = .73$. (The slopes of the P-center regression lines for the individual subjects are .95, 1.07, and .84; the correlations are .80, .91, and .57, respectively. Each correlation is significant at the .001 level). Thus, there is essentially a millisecond shift in P-center location for every millisecond of frication deleted.

There is no abrupt shift in P-center location at the category boundaries (which are indicated by the arrows). This indicates that the phonetic identity of syllable-initial consonantal segments does not affect P-center location. That there is no phonetic (or any other) source of nonlinearity in the data is revealed by a goodness-of-fit test. This test reveals that the first-degree polynomial is the highest one to significantly reduce the residual sum of squares. The significant F , $F(3,596) = 233.8$, $p < .001$, for degree 0 indicates that there is systematic variance unaccounted for at that level; the nonsignificant F for degree 1, $F(2,596) = .33$, n.s., indicates that the linear function accounts for as much of the variance as any higher polynomial. Thus, the point on the graph just after the category boundary does not signal a departure from linearity.

Experiment 2

The purpose of the second experiment was to replicate and extend the results of the first experiment using a /sa/-/sta/ continuum. In contrast to Experiment 1, the onset characteristics (including the rise time) of the stimuli in this experiment were left unaffected by the experimental manipulation. By eliminating the confounding effects of rise time, we are able to address Howell's (1984) claim that the amplitude envelope of a syllable (manipulated in his experiment by varying the syllable's rise time) significantly affects its P-center location.

Method

Stimuli. An eleven-step /sa/-/sta/ continuum was created by inserting 10-ms increments of silence between the frication and the vocalic segment of a naturally spoken /sa/ syllable. The frication of the original syllable was 206 ms in duration and the vocalic segment was 360 ms in duration. The first stimulus of the continuum was the original /sa/. The final stimulus contained a 100-ms gap between the frication and the vocalic segment (see Figure 4).

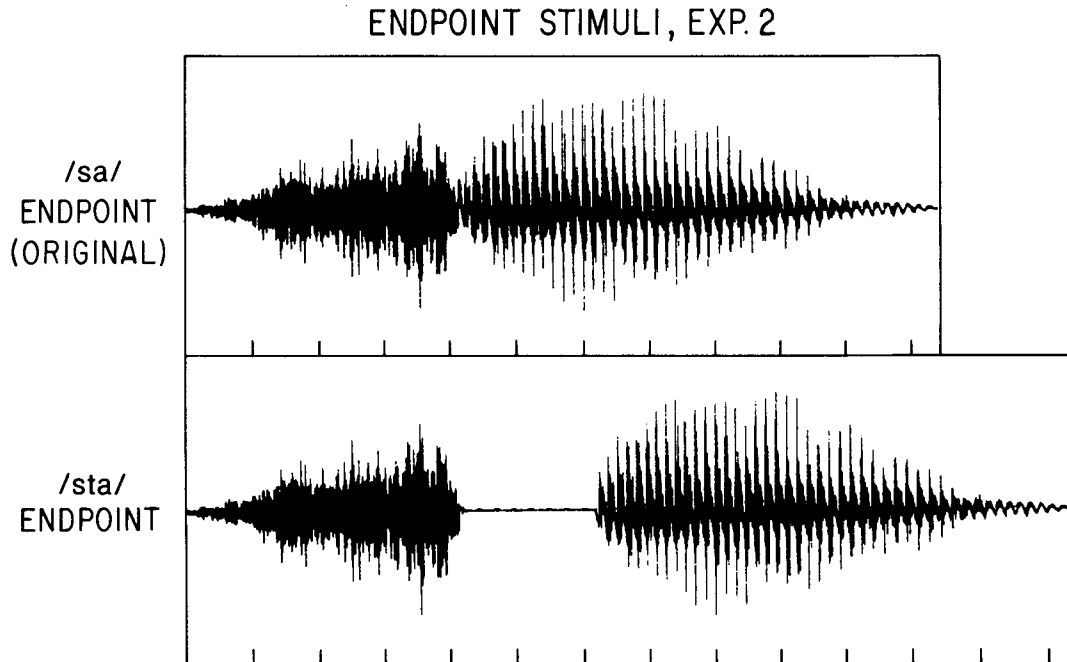


Figure 4. Continuum endpoint stimuli for Experiment 2: unmodified stimulus (upper panel), extreme stimulus manipulation (lower panel).

The identification test consisted of a randomized sequence of 20 repetitions of each member of the continuum (20 x 11 = 220 trials). A two-step AXB discrimination test consisted of a randomized sequence of four repetitions of the four versions of each of the nine pairings of the stimuli, within the AXB paradigm (9 x 4 x 4 = 144 trials). A two-step discrimination test was used rather than a one-step discrimination test (as in Experiment 1) because we found 10-ms differences between stimuli to be too small for listeners to discriminate consistently. (The stimuli in the one-step comparisons in Experiment 1 differed by 15 ms.) For the alignment test, 12 judgments were obtained for each member of the continuum (12 x 11 = 132 trials).

Subjects. Three subjects participated in the experiment. One subject was naive as to the purposes of the experiment. The other two subjects were two of the authors (AMC and CAF).

Results and Discussion

Figure 5 shows the mean identification and discrimination results for the three subjects. The ordinate represents the percentage of /sta/ responses for the identification data and percent correct for the discrimination data. The abscissa represents gap duration. The identification function shows an abrupt shift in phoneme identification. The mean category boundary occurs at about 60 ms of silence. The discrimination function shows a peak in discrimination near the mean category boundary and troughs in performance within categories. (The high discriminability of the first stimulus, when compared to the third, the point enclosed in parentheses, may have occurred because listeners were able to distinguish the unmodified stimulus from a stimulus that had been modified).

The mean results of the three subjects' performance on the alignment test are shown in Figure 6. The ordinate represents the displacement from acoustic isochrony of the test stimuli relative to the reference syllable in milliseconds. The abscissa represents the amount of silence inserted into the test stimuli in milliseconds.

In this experiment P-center location, as shown by the regression line, moves linearly toward stimulus offset across the continuum with a slope of 1.00, $r = .75$. (The slopes of the P-center regression lines for the individual subjects are 1.03, .94 and 1.04; the correlations are .82, .71 and .77. Each correlation is significant at the .001 level). Thus, as in Experiment 1, there is a 1-ms shift in P-center location for every millisecond of change in gap duration. The phonetic identity of syllable-initial consonantal segments does not affect the P-center; that is, there is no abrupt shift in P-center location at the category boundary (which is indicated by the arrow). A goodness-of-fit test provided an outcome analogous to that performed on the data from Experiment 1. Thus, there is no significant departure from linearity in the data.

The results of both Experiments 1 and 2 demonstrate a millisecond-for-millisecond shift in the P-center for each stimulus manipulation. Contrary to Howell's (1984) suggestion, in this experiment the P-center varied even though the rise time of the test stimuli remained constant. We have yet to determine, however, what aspect of the changing durational pattern accounts for our results. In particular, we want to know whether P-center shifts are a function of changes in gap size, of changes in overall stimulus duration, of changes in the duration of the prevocalic segment of the syllable, or of changes in the temporal location of the acoustically defined vowel: all of these change linearly and at the same rate in Experiments 1 and 2. The remaining experiments are designed to distinguish among these alternatives.

Experiment 3

Experiment 3 was designed to control for possible influences of variation in overall syllable duration on the P-center results of Experiment 2. Stimulus duration was held constant by excising an equal amount of the

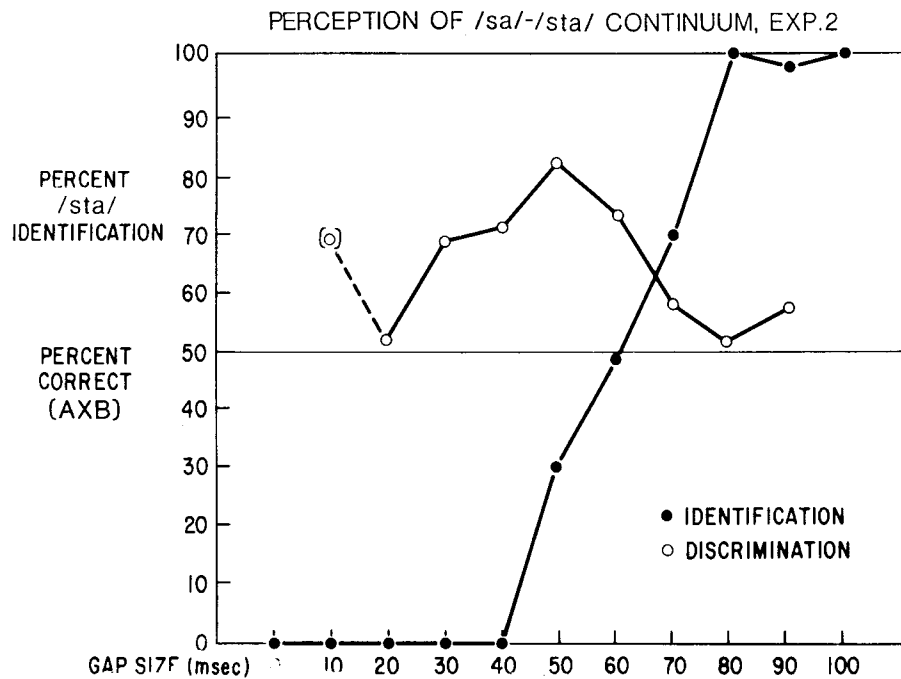


Figure 5. Mean identification and discrimination function (pooled across three subjects) for /sa/-/sta/ continuum, for Experiment 2.

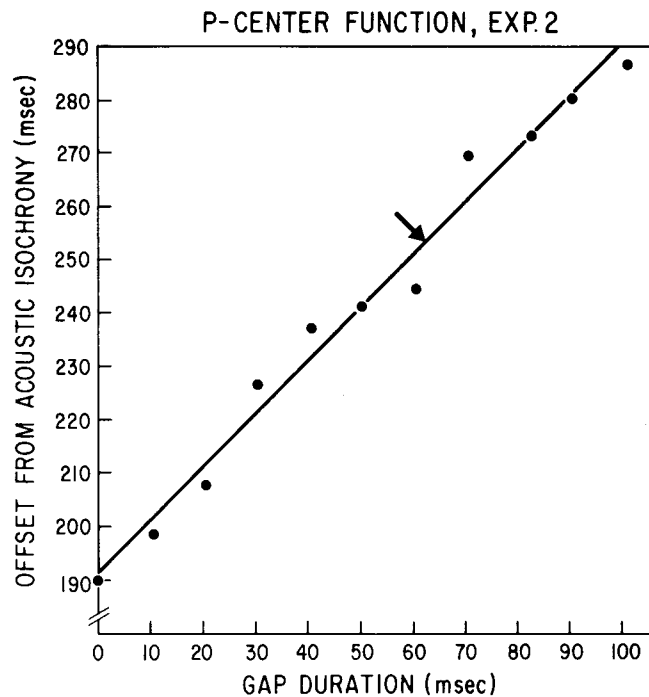


Figure 6. Mean P-center alignment function for Experiment 2, pooled across three subjects. The solid line represents the regression line and the arrow indicates the category boundary.

frication to offset the amount of silence inserted. If the shift in P-center noted in Experiment 2 was primarily related to gap duration, then the P-center functions of Experiments 2 and 3 should be equivalent. If, however, the observed shift in P-center is due either to prevocalic duration or to total duration of the syllable, then the P-center location should not change across the continuum.

Methods

Stimuli. The extreme stimuli of Experiment 3 are presented in Figure 7. The center waveform shows the unmodified /sa/ used in Experiments 2 - 4. In the final member of the continuum, represented by the upper waveform, 100 ms of silence has been inserted between the frication and the vocalic segment and 100 ms of frication has been deleted. For each stimulus in which silence was inserted, a compensatory amount of frication was excised beginning at a point 72.2 ms into the frication. This location was chosen because it allowed us both to maintain the original onset and offset characteristics of the frication and to excise a substantial amount of noise from within the fricative segment. The identification, discrimination, and alignment tests were organized as in Experiment 2.

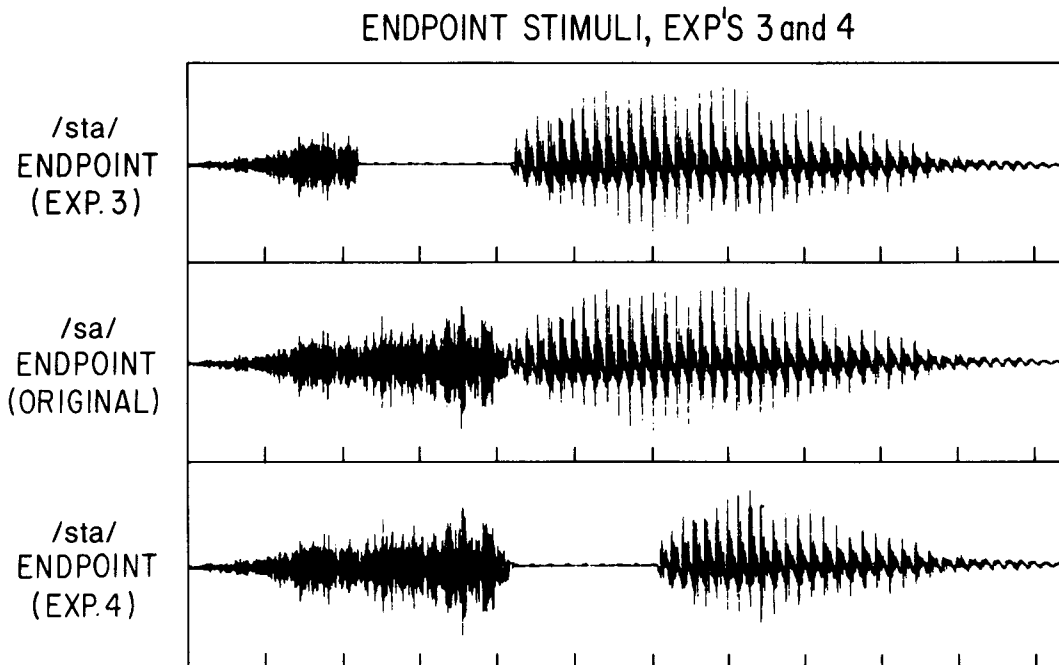


Figure 7. Continuum endpoint stimuli for Experiments 3 & 4: unmodified stimulus (center panel), extreme stimulus manipulation for Experiment 3 (upper panel), extreme stimulus manipulation for Experiment 4 (lower panel).

Subjects. The subjects were those of Experiment 2.

Results and Discussion

The identification data showed an abrupt shift in perceived phoneme category at 55 ms, and the discrimination data showed that discrimination is somewhat better near the category boundaries than within categories (Figure 8).

The alignment test for Experiment 3 (Figure 9) shows no significant change in P-center location across the continuum. The slope of the regression line is -0.003 ; the correlation is $= 0.004$. (The slopes of the P-center regression line for the individual subjects are -0.13 , 0.09 , and 0.02 ; the correlations are -0.14 , 0.18 , and 0.03 , respectively). This result shows that the linear shift in the P-center noted in Experiment 2 is not due to the increases in gap duration, per se. Nor, consistent with Experiments 1 and 2, is P-center location affected by the phonetic identity of syllable initial prevocalic segments. And finally, this experiment also shows that P-center shifts are not necessarily affected by a syllable's amplitude envelope (as suggested by Howell, 1984), since the envelopes of the stimuli varied although the P-center did not.

The canceling of the effect of the manipulation of gap duration on P-center location may be ascribed to the canceling of its effect on total syllable duration, to the canceling of its effect on the duration of the prevocalic consonant cluster, or to the canceling of its effect on the onset time of the vocalic segment. Experiment 4 is designed to distinguish among these alternatives.

Experiment 4

In Experiment 4, syllable duration was held constant, as in Experiment 3, however, but here compensation for increases in gap duration was achieved by shortening the vocalic segment of the syllable rather than the frication. If the variation in duration of the prevocalic segment or the onset time of the vocalic segment (acoustically defined) is principally responsible for the P-center shifts noted in Experiment 2, then the P-center functions of Experiments 2 and 4 should be equivalent. If total duration is responsible for the shifts in P-center, then the P-center should remain constant across the continuum (as in Experiment 3).

Method

Stimuli. The stimulus manipulations in Experiment 4 are illustrated in Figure 7. The center waveform represents the unmodified /sa/ used in Experiments 2-4. The lower waveform represents the final stimulus used in Experiment 4, in which 12 pitch periods were excised and 93 ms of silence were inserted. In the continuum, syllable duration was held constant by excising successive pitch pulses from the vocalic segment and then inserting compensatory amounts of silence. By excising complete pitch pulses from the vowel rather than excising an arbitrary amount of the vocalic segment, we avoided abrupt discontinuities in the periodic part of the signal. Pitch pulses were extracted beginning 67.7 ms into the vocalic segment. This location was both within the steady-state portion of vowel and beyond the peak intensity of the vowel. The individual pitch pulses ranged from 7.5 to 8.0 ms in duration. The identification, discrimination, and alignment tests were organized as in Experiment 2.

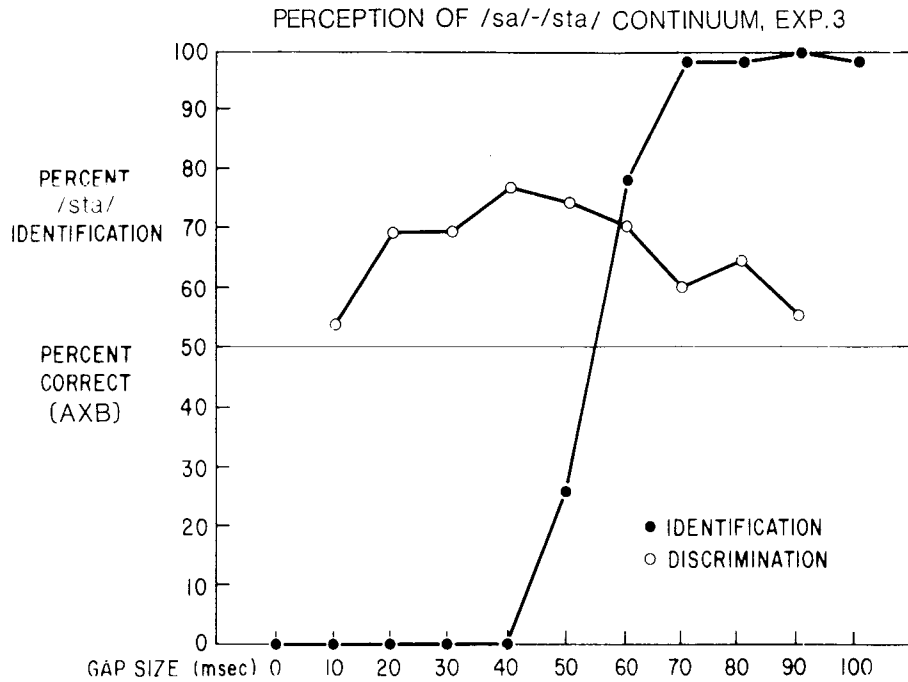


Figure 8. Mean identification and discrimination function (pooled across three subjects) for /sa/-/sta/ continuum, for Experiment 3.

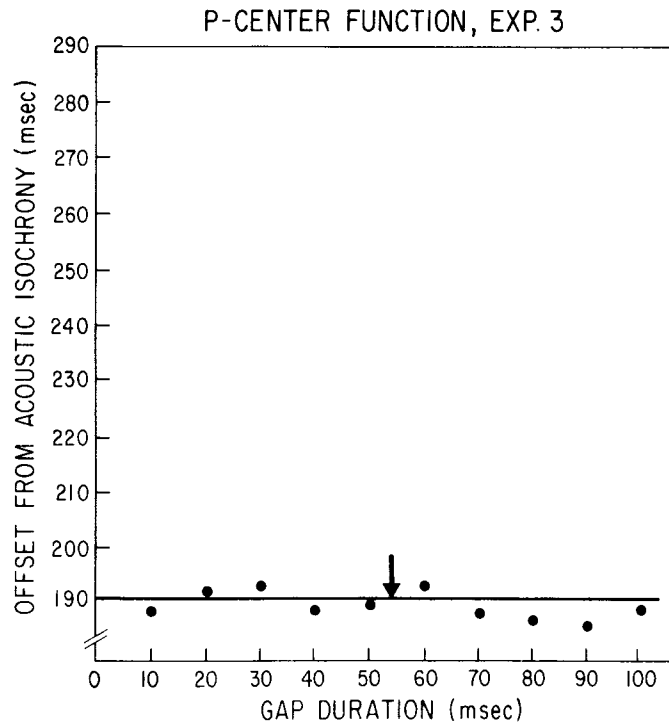


Figure 9. Mean P-center alignment function for Experiment 3, pooled across three subjects. The solid line represents the regression line and the arrow indicates the category boundary.

Subjects. The subjects were those of Experiments 2 and 3.

Results and Discussion

The identification data showed an abrupt shift in perceived phoneme category at 47 ms, and the discrimination data showed that discrimination was better near the category boundaries than within categories (Figure 10).

The results of the alignment test for Experiment 4 (Figure 11) show a linear shift in the P-center toward stimulus offset. The arrow indicates the category boundary. The slope of regression line is .83 with $r = .68$. (The slopes of the P-center regression lines for the individual subjects are .80, .80, and .88; the correlations are .88, .58, and .72, respectively. Each correlation is significant at the .001 level.) A dashed line with a slope of one is shown for comparison. A goodness-of-fit test reveals that the function is linear with no significant departure from linearity.

In this experiment, the P-center shifts toward stimulus offset, but the change is less than 1 ms of shift in the P-center for each millisecond of change in the gap duration. A paired t test comparing the slopes of the subjects' regression lines of Experiment 2 with those of Experiment 4 shows that these slopes are significantly different, $t(2) = 6.47$, $p < .05$. For the most extreme stimulus in Experiment 4, a gap size of 93 ms results in a P-center shift of 74.5 ms, whereas, in Experiment 2, a comparable gap size of 90 ms, shifts the P-center 91.4 ms--a difference of about 17 ms.

Our results are in agreement with those of Marcus (1981) who showed that P-center location is determined by the temporal makeup of the entire stimulus. We can interpret the results of Experiment 4 as follows. Increases in the duration of the prevocalic segment cause the P-center to shift toward stimulus offset, as it did in Experiments 1 and 2. Compensatory decreases in vowel duration, however, shift the P-center toward stimulus onset, although the magnitude of the shift is less. This interpretation is supported by Marcus (1981), who has shown that as the vocalic segment in CV syllables decreases in duration, the P-center shifts toward stimulus onset.

The results of Experiment 4 also show that P-center location cannot be simply associated with acoustically defined vowel onsets. For if the onset of the vowel were correlated with the P-center results in the previous experiments, then there should have been a millisecond shift in P-center location for every millisecond that the vowel onset shifted in the present experiment. Our results, however, clearly show that the correlation between P-center location and vowel onsets is not perfect and that P-centers cannot be linked exclusively to any single acoustically defined event in the speech signal. Other investigators have also shown that P-center location seems to be correlated with something other than a word's acoustically defined vowel onset (Allen, 1972; Fowler & Tassinary, 1981; Rapp, 1971). In Allen's study, subjects were required to tap "on the beat" of a specified syllable in a sentence. In Fowler and Tassinary's study, subjects were asked to produce rhyming nonsense syllables in time to a metronome. In Rapp's study, subjects were also asked to produce nonsense syllables in time with a regularly occurring pulse. In each of the studies, the pulse or tap both preceded the acoustic onset of the vowel and was positively correlated with the duration of the initial consonant or consonant cluster.

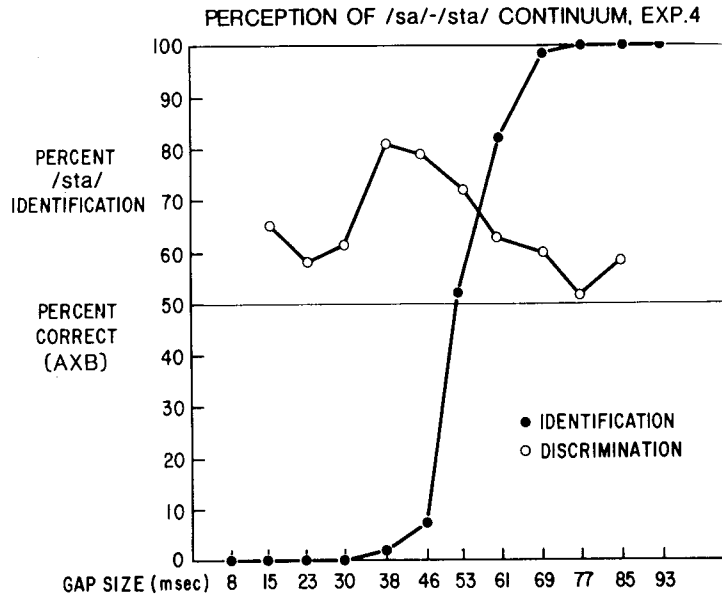


Figure 10. Mean identification and discrimination function (pooled across three subjects) for /sa-/sta/ continuum, for Experiment 4.

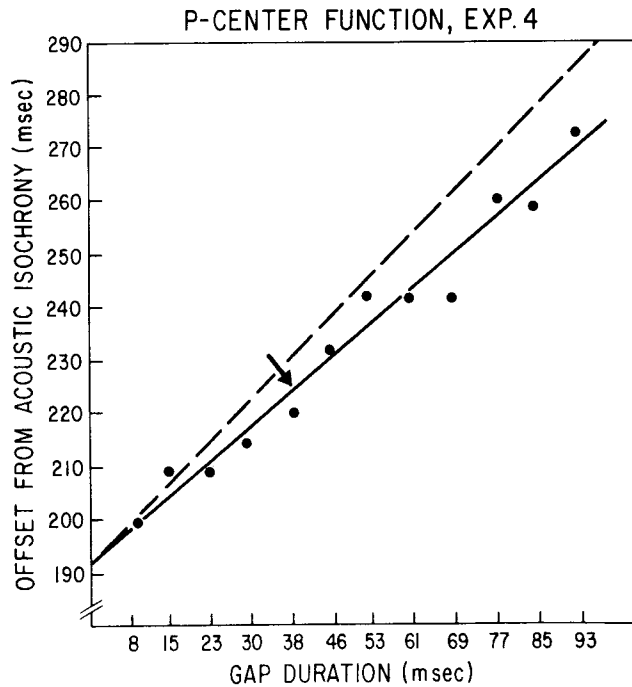


Figure 11. Mean P-center alignment function for Experiment 4, pooled across three subjects. The solid line represents the regression line, the broken line has a slope of 1 for comparison and the arrow indicates the category boundary.

Summary and General Conclusion

Our results demonstrate that a listener's judgments of the P-center of a syllable are not affected by the phonetic identity of the prevocalic segment or by any obvious acoustic properties of the signal, such as gap duration or simply the overall duration of the stimulus. Instead, the P-center appears to have been determined by a combination of at least two different aspects of the signal, the duration of the prevocalic segment and, to a lesser extent, the duration of the vocalic segment.

Our results also bear upon rationales for P-center shifts based on the amplitude envelope of speech stimuli. Howell (1984) performed experiments in which the onset intensity envelope of a /ʒa/ syllable was altered. He claimed that this manipulation was a "sufficient" source of variation in P-center location. Furthermore, he suggested that all of Marcus's manipulations (Marcus, 1981) in which the P-center was altered could be attributed to an alteration of the distribution of energy in the amplitude envelope. Although temporal judgments of nonspeech stimuli are influenced by their rise time characteristics (Howell, 1984; Vos & Rasch, 1981), this explanation does not account for P-center shifts in speech stimuli. For example, when the onset envelope of the test stimuli remained unaltered (Experiments 2 and 4), the P-center location varied; in contrast, when the onset envelopes of the test stimuli were altered, (Experiment 3), the P-center did not vary in location. Consistent with this finding, Tuller and Fowler (1980) radically changed the amplitude envelope of speech syllables by infinite peak clipping and found no shift in the P-center. Finally, Marcus (1981) found that increases in the silent interval for the dental stop in the word "eight" shifted the P-center toward the acoustic offset of the word, but that increases in the amplitude of the release burst did not affect the location of the P-center.

Although our findings show that the phonetic identity of syllable-initial consonants does not affect the location of the P-center, they do not rule out any possible effect of the phonetic structure of a syllable on P-center location. Given that P-center location shows a precise millisecond-for-millisecond relationship with the duration of prevocalic segments in a syllable, but is affected to a markedly smaller extent by the duration of the vocalic segment, the results of Experiment 3 are open to both phonetic and acoustic interpretations. One interpretation is that this diminished durational effect on P-center location occurs abruptly just at the point in a syllable where the listener's perception of prevocalic segments gives way to perception of the vowel. We would identify this effect as phonetic and, accordingly, we would predict that manipulations of vowel duration in vowel-initial syllables would have effects on P-center location equivalent to the vocalic effects in Experiment 4. An alternative explanation is that the effects of changes in duration are weaker the farther away they are from the syllable onset. We would identify such an effect as acoustic, and would expect the effects of durational manipulations of vowels to be greater in vowel-initial syllables than in those with initial consonants. We are currently investigating whether the phonetic quality of syllable constituents as well as the serial position of those constituents within a syllable affect the location of the P-center.

References

- Allen, G. (1972). The location of rhythmic stress beats in English: An experimental study, I. Language and Speech, 15, 72-100.
- Best, C. T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. Perception & Psychophysics, 29, 191-211.
- Fowler, C. A., & Tassinari, L. G. (1981). Natural measurement criteria for speech: The anisochrony illusion. In J. Long & A. Baddeley (Eds.), Attention and Performance (Vol IX, pp. 521-535). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Healy, A. F., & Repp, B. H. (1982). Context sensitivity and phonetic mediation in categorical perception. Journal of Experimental Psychology: Human Perception and Performance, 8, 68-80.
- Howell, P. (1984). An acoustic determinant of perceived and produced anisochrony. In M. P. R. Van den Broecke & A. Cohen (Eds.), Proceedings of the 10th International Congress of Phonetic Sciences (pp. 429-433). Dordrecht: Foris Publications.
- Lieberman, A. M., Harris, K. S., Eimas, P. D., Lisker, L., & Bastian, J. (1961). An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance. Language and Speech, 4, 175-195.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 54, 358-368.
- Lieberman, A. M., Harris, K. S., Kinney, J. A., & Lane, H. (1961). The discrimination of relative onset time of the components of certain speech and nonspeech patterns. Journal of Experimental Psychology, 61, 379-388.
- Marcus, S. M. (1976). Perceptual centres. Unpublished doctoral dissertation, Cambridge University.
- Marcus, S. (1981). Acoustic determinants of perceptual center (P-center) location. Perception & Psychophysics, 30, 247-256.
- Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). Psychological Review, 83, 405-408.
- Rapp, K. (1971). A study of syllable-timing (Papers in Linguistics). Stockholm: University of Stockholm.
- Repp, B. H. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), Speech and language: Advances in basic research and practice (Vol. 10). Orlando, FL: Academic Press.
- Tuller, B., & Fowler, C. A. (1980). Some articulatory correlates of perceptual isochrony. Perception & Psychophysics, 27, 277-283.
- Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. Perception & Psychophysics, 29, 323-335. 68-80.