

# Secant lines in a preliminary study of phonetic reduction of /ɪ/ and /ɛ/

Richard S. McGowan<sup>a)</sup>

*CReSS LLC, 1 Seaborn Place, Lexington, Massachusetts 02420-2002*

(Received 8 May 2003; revised 13 May 2004; accepted 13 July 2004)

Two subjects from the X-Ray Microbeam Speech Production Database were examined in their production of the vowels /ɪ/ and /ɛ/ in alveolar and dental consonant contexts. Secant lines, or first-order splines, between the three most anterior pellets were examined at vowel critical times. These critical times were zero crossings in the tangential acceleration of the midpoints of the secant lines. We expected and found, in general, that vowel reduction occurred as a function of vowel duration in measures of the secant line midpoint-to-palate distance and secant line orientation at vowel critical times. The shorter the vowel, the smaller the distance of the secant line midpoints to the palate and the less downward the orientation of the secant lines at the vowel critical times. Phonetic reduction was also apparent in the formant frequencies. There were differences between the speakers in terms of the range of vowel duration and degree of reduction. The subjects differed in the functional parts of the tongue spanned by the secant lines and the shape of their palates. These differences were factors in the observed relations between formant frequencies and the articulatory, secant line measures for each subject. © 2004 Acoustical Society of America.

[DOI: 10.1121/1.1789491]

PACS numbers: 43.70.Aj [AL]

Pages: 2324–2337

## I. INTRODUCTION

### A. Previous work

Tongue kinematics in consonant–vowel and vowel–consonant transitions needs to be understood in order to perform articulatory synthesis of English and other languages. One particular phenomenon associated with these transitions is phonetic vowel reduction. Phonetic vowel reduction in syllables is a complex, speaker-dependent phenomenon that is often characterized by greater coarticulation of the vowel with the surrounding consonants as the duration of the vowel decreases. In this paper, we seek to understand some articulatory and acoustic aspects of phonetic vowel reduction in English front vowel production in terms of tongue positions and shapes and their consequent acoustics.

Lindblom (1963) quantified reduction of Swedish lax vowels in the acoustic domain by considering formant frequencies of the vowels as a function of vowel duration for one speaker. He showed that the shorter duration vowels tended to possess formant frequencies at their centers that were more greatly affected by the surrounding segmental context than did those of longer duration. Thus, he concluded that vowel reduction in languages with “heavy stress,” like Swedish and English, is a phenomenon of assimilation and not of neutralization. Lindblom (1964) verified what he observed in his 1963 study on another Swedish speaker with both tense and lax vowels. Stevens and House (1963) also showed that the first two formants in American English vowels spoken in words with symmetric consonant context were affected by the places of articulation of those consonants. Again, the vowels showed assimilation, or coarticulation, with the surrounding consonants.

There have been a number of acoustic studies of coarticulation in vowel production since the original Lindblom and Stevens and House papers appeared. Stevens, House, and Paul (1966) noted asymmetries in second formant frequency, or F<sub>2</sub>, trajectories for vowels in symmetric consonant context of nonsense syllables. Further, for lax vowels, there was a tendency for F<sub>2</sub> to approach a neutral value at closure of the final consonant. Ohde and Scharf (1975) observed that the onset consonant had more influence on the succeeding vowel than did the vowel’s succeeding consonant. The Stevens *et al.* (1966) and Ohde and Scharf (1975) observations could indicate both a coarticulatory effect of surrounding consonants on the vowel, and, perhaps, a certain degree of lenition of the final consonant, particularly when it is preceded by a lax vowel. Broad and Fertig (1970) were able to show that a linear superposition of the consonant effects on the vowel /ɪ/ in nonsymmetrical CVC syllables provided a good statistical model of formant trajectories. Broad and Clermont (1987) refined this model to include more vowels and to obtain particular formant trajectory shapes.

Acoustic investigations have progressed substantially in considerations of factors other than vowel duration in vowel reduction. Gay (1978) found no evidence of vowel reduction with changes in speaking rate and stress. However, this could have been the result of using vowels in voiceless, bilabial consonant context, and the result of instructions given to the speakers. Nord (1986) employed acoustic analyses to show that factors other than duration, such as phrase position and stress, have an effect on the degree of vowel reduction. He studied vowels in syllables with alveolar/dental consonant onsets and codas, and these syllables were parts of two syl-

<sup>a)</sup>Electronic mail: rsmcgowan@earthlink.net

lable words. Di Benedetto (1989) studied the properties of F1 trajectories in three English speakers and found further evidence of assimilation with the surrounding consonants. She also showed that F1 and F2 sampled at maximum F1 alone could not be used to distinguish between /t/ and /ɛ/. However, by considering other F1 trajectory properties, such as the proportion of total vowel duration when the maximum is attained, these vowels could be distinguished. Van Son and Pols (1990, 1992), in studying the read speech of a professional newscaster, found minimal effects on formant tracks when vowel durations were changed as a result of speech rate changes. The only consistent change that they found was that of a higher first formant frequency for all vowels in the fast speaking rate condition. This change in F1 could have been the result of a consistently more open mouth, or other postural change in the vocal tract, during faster rate speech. They concluded that the speaker changed speaking style with speaking rate in order to maintain midvowel formant frequencies.

Further work has confirmed that speaking style and suprasegmental conditions are major factors in vowel reduction phenomena. van Bergem (1993) performed a study on the effects of word stress, word class, and sentence stress on the vowels in Dutch read sentences. He found that word stress and word class were the major determiners of F1 and F2. Further, he confirmed that vowel reduction was a consonant assimilation phenomenon and not the result of neutralization. van Bergem concluded that vowel reduction is as much a matter of speaking style as it is a matter of vowel duration. Moon and Lindblom (1994) also studied vowel reduction using formant frequency data. They asked speakers to produce speech in clear and citation-form styles at a constant speaking rate. They found the expected assimilation of vowel formants with surrounding segments, but that clear speech involved an “undershoot-compensating” mechanism. By adding the extra effect of speaking style, they extended the results of Lindblom (1963).

The intricacies of vowel reduction are also apparent in articulatory studies of tongue movement. Lindblom (1964) confirmed that, for a single speaker producing a single consonant-vowel combination, articulators moved a smaller distance from consonant release into the vowel when the duration of the vowel was diminished. Kuehn and Moll (1976) noted that simple instructions to talkers to increase the rate of speech of nonsense words and syllables sometimes resulted in either increased articulatory velocity, decreased articulatory displacement, or a combination of these effects. Flege (1988), who studied rate-induced vowel reduction in schwa and bilabial consonant environments, noted some vowel reduction, as well as production strategies that mitigated vowel reduction. These strategies included increased peak tongue velocities, as well as early movement toward the stressed vowel and preservation of the vowel in movements away from the vowel. Engstrand (1988) experimented with Swedish tense vowels, also in schwa and bilabial consonant contexts. He noted that the stressed vowels were more extreme in the formant frequencies and they were produced with tighter vocal tract constrictions. He found little effect of rate on vowel reduction, but noted that his

results could have been different in other consonantal contexts. More recently Adams, Weismer, and Kent (1993) have confirmed and extended the observations of Kuehn and Moll (1976) that speakers have a variety of articulatory strategies when changing speaking rate. Finally, Perkell, Zandipour, Matthies, and Lane (2002) have shown that speakers do use different articulatory means to change from normal style to clear style speech.

## B. The present study

The present study seeks to supplement and extend the previous findings on changes in vowel articulation and acoustics with vowel duration. In particular, front vowels /t/ and /ɛ/ in alveolar and dental consonant environments for two English speakers were examined using *secant lines*, which are described below. The data were from the Wisconsin X-Ray Microbeam Speech Production Database, the XRMB-SPD (Westbury, 1994), which contains digitized tongue flesh point trajectories for adults producing various word lists, sentences, and paragraphs.

There are many factors that determine the duration of the vowel, such as rate, degree of emphasis or stress, phrase position, and the idiolect of individual speakers. We chose to examine the effect of differences in vowel duration within individual speakers, without regard to the linguistic or non-linguistic causes of these differences. Part of this was driven by a desire to see whether the degree of vowel reduction or consonant assimilation effects could be detected as a function of vowel duration alone, absent reference to the causes of the vowels' absolute durations. The other reason is that the data could not be readily separated according to rate, stress, and phrase position without reducing the number of tokens within each category to the point that statistically significant trends could be detected.

This study of vowel reduction is a part of a larger project to base articulatory synthesis on kinematic data. More fundamentally, the goal is to understand and verify models of the articulatory-acoustic relations. With the most readily available kinematic data in the form of flesh point data as collected by the x-ray microbeam system (Fujimura, Kiritani, and Ishida, 1973; Westbury, 1994) and the ElectroMagnetic Midsagittal Articulometry, EMMA (Perkell, Cohen, Svirsky, Matthies, Garabieta, and Jackson, 1992), it is important to be able to analyze these data employing articulatory representations that support these goals. Primarily, the articulatory representations of flesh point data needs to provide tongue shape information, because the overall shape of the tongue surface determines the vocal-tract area function, and, thus, the resonance properties of the vocal tract. The line segments between neighboring flesh points are called *secant lines* here. Secant lines provide a first step in characterizing the geometric relations between the tongue surface and the fixed structures, such as the hard palate, for flesh point data. Although secant line kinematics can provide only a first approximation to tongue shape kinematics, it does so without the much more complicated continuum description. Further, it is a local description in that changes in one secant line do not change a secant line that does not share a pellet. This is unlike PCA or RARAFAC analyses (e.g., Hoole, 1999),

which characterize statistical regularities in the global shape of the tongue by a small number of basic global shapes. The immediate motivation here was to find how vowel reduction phenomena could be characterized within a secant line representation. In this work we were concerned with following the movement of two of the most anterior secant lines between the three most anterior pellets on the tongues of two subjects.

The procedure of tracking secant line kinematics has been foreshadowed by Stone (1990), who studied CV and VC transitions in read speech with the x-ray microbeam and ultrasound, where the consonants were /l/ and /s/ and the vowels were /i/, /a/, and /o/. She concluded: “*The tongue can be segmented sagittally into functional segments that operate as semi-independent units within the tongue. Movements of these segments result in local displacement and anterior to posterior tongue rotation*” (Stone, 1990; p. 2216). The main evidence for this statement came from two kinds of measurement. One was that the change in the distance between adjacent pellets among various phonemes depended on which part of the tongue the pellets were attached and on the phoneme type. For instance, the distance between the third and fourth of five pellets always seemed to be smaller than for the others in Stone’s test utterances, despite the fact they were set on a protruded tongue at equal distances. The other evidence came from the relative timing of articulatory events. The anterior-most or the posterior-most pellets reached their extreme positions for neighboring phonemes in the same temporal order. The differing inter-pellet lengths and relative timings indicated that the sections of the tongue were controlled somewhat independently. Stone hypothesized five independent sagittal segments in all from the tongue tip to the tongue root. Further, Westbury, Hashi, and Lindstrom (1998) have used secant lines to characterize the tongue shapes of various tokens of [ɪ] in the XRMB-SPD.

### C. Secant line kinematics

Typical tongue flesh point data contain time-sampled positions of two to five pellets placed on the midline of the tongue from just behind the tongue tip to, perhaps, 6 cm or so behind the tip. The XRMB-SPD usually has records of four pellets on the midline of the tongue surface. The motion of four midsagittal flesh points on the tongue can be characterized by the motion of each of two secant lines: one between the two most anterior flesh points and another between the two most posterior flesh points. In this paper we were most concerned with the kinematics of the anterior of the tongue during front vowel production. Thus, we considered the kinematics of the two most anterior secant lines, which involved the three most anterior pellets. Numbering the pellets from anterior to posterior, we considered the secant lines between the first and second pellets and the secant line between the second and the third pellets. The former secant line is referred to as the 1–2 *secant line*, and the latter secant line is referred to as the 2–3 *secant line* (see Fig. 1). Each secant line possesses four degrees of freedom in the midsagittal plane: two translation degrees of freedom of the midpoint of the secant line, the rotation about the midpoint, and the dila-

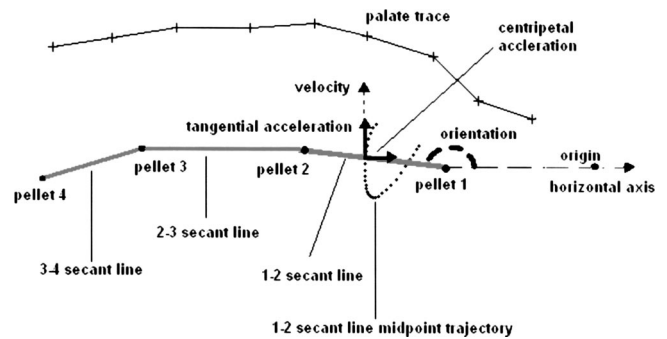


FIG. 1. The 1–2, 2–3, and 3–4 tongue secant lines for JW29. The dashed line with the arrow at the end of pellet 1 is the  $x$  axis (horizontal direction). The palate trace is also shown. The details of the 1–2 secant line kinematics are shown. The locus of dots, or points, represents the time-sampled path of the midpoint of the 1–2 secant line through a vowel: one for each time sample at 160 Hz. At any particular time, the velocity and tangential acceleration vectors are parallel, and the centripetal acceleration is orthogonal to these. The orientation of the 1–2 secant line is the angle between the horizontal direction and the 1–2 secant line.

tion, or change in length, of the secant line. We studied the changes in translation and rotational degrees-of-freedom as functions of vowel duration in this work.

The kinematic quantities defined below are illustrated in Fig. 1 for the 1–2 secant line. The midpoint of each secant line travels along a path, or locus of points, in the midsagittal plane. The time derivative of the midpoint position vector is the velocity of the midpoint. This is a two-dimensional vector with a direction that specifies the *tangent direction* and a magnitude, or *speed*, at each point on the path. The time derivative of the velocity can also be calculated to obtain acceleration. Because the velocity has two spatial degrees of freedom, the acceleration has two spatial degrees of freedom. One of these specifies the time rate of change of the speed in the tangent direction, and the other specifies the time rate of change of the tangent direction itself. The first component is parallel to the tangent direction and is known as *tangential acceleration*. While translation speeds are positive by definition, tangential acceleration can either be negative or positive, because the translation speed can either be increasing or decreasing. The second component of acceleration is directed orthogonal to the tangent and is known as *centripetal acceleration*. Another part of the specification for the secant line is in terms of *orientation angle* with respect to the horizontal direction. The orientation angle increases when the secant line rotates about its midpoint in the counterclockwise direction and the orientation angle decreases when the rotation is clockwise. In the former case the anterior end of the secant line points more upward, and in the latter case it points more downward. The origin of the coordinate system is that defined in the XRMB-SPD handbook (Westbury, 1994, p. 37), and is near the tips of the upper incisors in the midsagittal plane. The  $x$ -axis is formed by the intersection of the midsagittal plane and a plane that intersected the tips of the central incisors and at least two other maxillary teeth on opposite sides of the mouth. The  $y$  axis is the line in the midsagittal plane, passing through the origin that is orthogonal to the  $x$  axis.

TABLE I. Speech tokens for subjects JW11 and JW29.

Word	Context	Number of tokens for JW11	Range of vowel durations for JW11	Number of tokens for JW29	Range of vowel durations for JW29
this	Word list	5	163–225 ms	5	225–256 ms
	“...damage <b>this</b> ship’s hull.”	1	100 ms	1	106 ms
	“Put <b>this</b> one right here.”	1	94 ms	1	88 ms
	“But <b>this</b> March,....”	0		1	100 ms
	“ <b>This</b> March, Tom...”	0		1	88 ms
sense	Word list	5	106–150 ms	5	163–181 ms
	“...provide a <b>sense</b> of order.”	4	106–131 ms	5	100–125 ms
	“...make <b>sense</b> of the problem.”	5	81–113 ms	5	100–118 ms
said	Word list	1	194 ms	1	250 ms
	“They all know what I <b>said</b> .”	5	163–194 ms	5	206–238 ms

## II. METHOD

Two subjects, JW11 (male) and JW29 (female), were chosen from the XRMB-SPD for examination. These subjects were chosen based on gender and because they had different palate shapes, and part of the inquiry was to address differences caused by palate shape. There turned out to be other differences that appeared more obvious as the data were analyzed: the two speakers had different articulatory behaviors as a function of vowel duration and corresponding pellets did not mark the same functional parts of the tongue. The XRMB-SPD contains articulatory data from several channels and acoustic data that have been synchronized using mathematical interpolation. All articulatory data channels have the common frame rate of 160 samples/s, which means that the duration between frames is 6.25 ms (Westbury, 1994, p. 57). The output speech was sampled at 21379 samples/s with a microphone placed 10 cm anterior and 5 cm lateral to the mouth opening (Westbury, 1994, p. 29).

We noted those words that contained either the front vowel /i/ or /e/ for subjects JW11 and JW29, and that were immediately preceded and followed by alveolar or dental consonants or consonant clusters. The words that met these criteria, and for which there were more than one or two tokens, were “this,” “sense,” and “said.” Thus, the initial consonants considered were /s/ and /ð/, and the final consonants or consonant clusters were /d/, /ns/, and /s/. These words appeared in various contexts: word lists, lists of sentences, and stories. Table I gives the numbers of tokens, contexts, and range of vowel durations for JW11 and JW29. The times of release of the initial consonant and closure of the final consonant or first consonant of a final consonant cluster were marked based on the acoustic record. The times of release of the fricatives were taken at the cessation of frication. Closures were marked by abrupt changes in acoustic amplitude in frequencies above 1 kHz for /d/ and /ns/ and by the beginning of frication for /s/. These were marked to the nearest articulatory sample time, rounding down to the preceding articulatory sample time for releases and up to the following articulatory sample time for closures. The time interval be-

tween each release and the closure was the vowel duration.

We chose to analyze the 1–2 and 2–3 secant lines. Without constraints on the movement of the tongue, movements of the 1–2 or 2–3 secant lines cannot be used to infer what happens in posterior regions of the tongue. For example, the translational movement of either or both the 1–2 or 2–3 secant lines could be caused by a general downward translation of the tongue body or a rotation of a portion of the tongue in the oral cavity with an axis of rotation posterior to the midpoints of the secant lines in question, as could be caused by jaw rotation. In the data examined here, we argue that differences in the acoustics caused by the parts of the tongue posterior to the 1–2 and 2–3 secant lines is minimal within a vowel, because of bracing of the sides of the tongue on the palate and teeth for mid to high front vowels (Perkell, 1991). This physical constraint should be further enhanced because these vowels appear in coronal consonant contexts. Also, not only would jaw motion be limited, but also its effect on the tongue should be limited by the bracing and by the fact that the tongue is not rigidly connected to the jaw.

For subjects in the XRMB-SPD, four pellets were placed along the midline (longitudinal sulcus) of the tongue, so that the most anterior pellet was about 1 cm posterior to the apex (tongue tip) and the most posterior about 6 cm from the apex, as long as a gag reflex was not encountered. The two other pellets were approximately evenly spaced between these extremes, and there were no gender-related differences in pellet spacing and placement (Westbury, 1994). Pellet placement can differ between subjects in the XRMB-SPD. This issue can be addressed by briefly considering the distances between the pellets for our two subjects in sustained vowels /a/ and /ʌ/, because these vowels were produced with a fairly flat tongue blade. For JW11 the average distance between pellets one and two was 16.2 mm, and between pellets two and three the average distance was 14.4 mm. For JW29 the average distance between pellets one and two was 17.3 mm, and between pellets two and three the average distance was 17.7 mm. Thus, the first three pellets had a greater separation for JW29 than for JW11. On the other

hand, the vocal tract appeared to be larger for JW11, an adult male, than for JW29, an adult female. The midsagittal palate outlines of the two subjects indicated a horizontal span of about 60 mm for JW11 and 50 mm for JW29. The line that was meant to be a rough indication of the position of the posterior pharyngeal wall was about 88 mm behind the anterior of the palate trace for JW11 and about 75 mm behind the anterior of the palate trace for JW29. These data indicate a smaller vocal tract, and probably, a smaller tongue for JW29 than for JW11. Further, the palate of JW11 was more domed than that of JW29 [see Figs. 4(a) and 4(b) later for palate shapes], which is yet another factor that makes the effective acoustic path length still longer for JW11 than for JW29. These factors, together with the fact that the pellets were more widely spaced for JW29 than for JW11, indicates that a functionally more posterior part of the tongue was considered for JW29 than for JW11 when examining the positions and orientations of the 1–2 and 2–3 secant lines.

An analysis program was written to calculate kinematic quantities for the 1–2 secant lines and the 2–3 secant lines. Midpoint velocities were calculated numerically using centered differences for each of the horizontal and vertical components of the secant line midpoints. The centered-difference for a digitized signal,  $s(t)$ , at sample time  $t_i$ , where  $i$  is a sample index, is  $(s(t_{i+1}) - s(t_{i-1})) / (t_{i+1} - t_{i-1})$ . Speed is the magnitude of the velocity vector. Tangential accelerations were computed numerically using centered differences of the speeds.

In the analysis, we were concerned with characterizing tongue position and shape at consonant release, closure, and in the middle of the vowel. (The choice of time for “middle of vowel,” denoted here as *vowel critical time*, is discussed later in Sec. III.) The articulatory measures used were the distance of the midpoint of each secant line to the palate and the orientation of each secant line with respect to the horizontal direction. The former measure is called *secant line midpoint-to-palate distance*, and the latter quantity is called *secant line orientation*. Four or five line segments approximated the midsagittal palate outlines for the purpose of calculating the distances between the secant lines and the palate. The midpoint-to-palate distance was calculated by calculating the lengths of the perpendicular from each palate line to the secant line midpoint. If the resulting perpendicular intersected its palate line outside of the range for which it represented the palate, the length of the line from the closest end point on the palate to the secant line midpoint was taken to be the distance to that portion of the palate. The minimum of the distances between each palate segment and the secant line midpoint was taken to be the midpoint-to-palate distance. We avoided discontinuities in this quantity by using such a calculation. An alternative strategy for calculating distances to the palate has been proposed by Lindstrom and Westbury. They calculate the minimum distance to the palate using an exponentially weighted average distance from each of the sampled palate points to avoid abrupt changes in this quantity with small changes in tongue position (Simpson, 2002).

Based on Lindblom’s 1963 analyses, it was expected that the differences between the measures at release or clo-

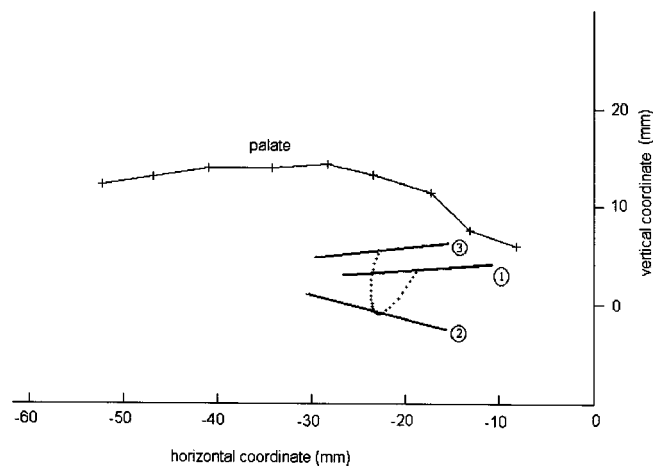


FIG. 2. Examples of a time sample’s 1–2 secant line midpoint trajectory as a dotted path for /ε/ in the word “sense” spoken by JW29. The dark line segments are the 1–2 secant line at (1) the release of /s/, (2) at a time when the movement goes from opening to closing, and (3) at the closure of /n/.

sure and at vowel centers could show a clearer trend with vowel duration than the absolute measures at centers alone, because vowel reduction is a coarticulation, or assimilation, phenomenon. Reference can be made to the 1–2 secant line and its time-sampled midpoint path for /ε/ in the word “sense” spoken by JW29 illustrated in Fig. 2. Three snapshots of the 1–2 secant line are shown: (1) at release, (2) at the maximum midpoint-to-palate distance, and (3) at closure. We can expect that the positions of the secant lines at release and closure will be mutually dependent on its position at vowel center. [Interestingly, consonant *jaw* position is affected more than vowel *jaw* position in English (Keating, Lindblom, Lubker, and Kreiman, 1994). We do not know of a comparable study for tongues.] Thus, not only were distances and orientations recorded in the vowel center, but the differences with the corresponding values at release and closure were also calculated.

The errors in articulatory measurement are discussed in the manual that accompanies the CDs containing the XRMB-SPD, Westbury (1994). From the manual, we expect the error in measuring the position of a moving pellet to be less than 1 mm. Thus, the error in measuring the position of the midpoint of a secant line will be less than 1 mm. The error in measuring the orientation of a secant line will be approximately, at most, 6°. The microbeam data were low-passed filter using spline fitting that is equivalent to fourth-order, zero delay Butterworth filtering. Corner frequencies for which the attenuation was 6 dB down from full pass were 12 Hz for the most anterior tongue pellet and 10 Hz for the other tongue pellets (Milenkovic, 2003). There were unequal numbers of tokens for each word extracted for each subject most often because of tracking problems during the recording of particular tokens.

These data allowed an exploration of some of the physics of the articulation-acoustic relationship. Formant frequencies were measured using 256-point windows to produce DFT cross sections at vowel centers. Formant frequency measurement errors made by the author were estimated using the measurements taken by another speech scientist who was

uninformed as to the purpose of this study. The average and standard deviation in the absolute value of the percent difference in each of the measured values by subject and formant frequency number were calculated. For JW11's F1s the average was 4% with a standard deviation of 4% and the average for his F2s was 2% with 1% standard deviation. For JW29's F1 the average was 7% with a standard deviation of 3% and the average for her F2s was 3% with 3% standard deviation.

### III. RESULTS

#### A. Qualitative patterns and vowel critical times

Comparisons of vowel durations between JW11 and JW29 indicate that JW11 did not distinguish contexts in which the tokens were spoken as much as JW29 did. For the vowel /i/ the ratio of the maximum to minimum vowel duration was 2.4 for JW11 and 3.7 for JW29. The range of durations for the vowel /ε/ produced a ratio of maximum to minimum duration of 2.4 for JW11 and 2.7 for JW29 (see Table I). The smaller ratios for JW11 indicated that JW11 and JW29 could be using different articulatory strategies in producing their vowels. In order to test the hypothesis that JW11 used articulatory compensation to mitigate vowel reduction, regression analyses of maximum 1–2 secant line midpoint speeds in the opening phase for the two subjects and two vowels were performed. A significant negative slope ( $p < 0.1$ ) occurred only for the case of JW11 producing /i/. Thus, JW11 sped his articulations up when the /i/ in “this” became shorter. Other investigators have noted changes in articulatory behavior that diminish vowel reduction effects (e.g., Kuehn and Moll, 1976; Harris, 1978; Ostry and Munhall, 1985). Speakers will often increase articulatory velocities or the force with which a consonant is released to mitigate the assimilation of the vowel to the surrounding consonants (Moon and Lindblom, 1994).

Because the vowels of interest were front vowels and the surrounding consonants were coronal, there were consistencies in secant line kinematics among the word tokens. This was particularly true for vowels of longer or moderate duration. Consideration of these consistencies allowed for a reasonable choice of criteria for vowel critical times (i.e., vowel centers) for vowels of at least moderate duration. Some of these criteria turned out to be useful for, and could be generalized to, vowels of the shortest duration that did not exhibit these consistencies. The consistencies that are being referred to are illustrated in Fig. 2. The 1–2 secant line midpoint generally had a downward movement away from the palate after the release of the initial consonant, with subsequent upward movement before the closure of the final consonant. Before the end of the downward movement, the midpoint slowed, and at the beginning of the upward movement the speed increased. (The closer the points on the path illustrated in Fig. 2, the slower the midpoint movement.) Thus, the maximum midpoint-to-palate distance often corresponded closely to a speed minimum and a zero crossing in tangential acceleration. Further, the general pattern of rotation for the 1–2 secant line was clockwise or decreasing orientation angle, after consonant release, as illustrated by

the snapshots of the secant lines at times (1) and (2) in Fig. 2. Also, the rotation tended to be in the opposite (counterclockwise) direction from sometime in the mid-portion of the vowel to the following consonant closure, as illustrated by the snapshots of the secant lines at times (2) and (3) in Fig. 2. This meant that the tip of the tongue tended to point downward more in the vowel than at the release or closure of the surrounding consonants in many, but not all, cases.

For the pattern just described, the center of the vowel was considered to be sometime during the change in the midpoint direction from downward to upward. Various criteria were considered for choosing times within this time interval to record quantities of interest: the maximum midpoint-to-palate distance, local midpoint speed minima, or zero crossings in the tangential acceleration. In some of the shorter utterances, there were no local speed minima, or they were very close to release or closure, the maximum midpoint-to-palate distance occurred near the release or closure, and/or the only zero crossing in tangential acceleration corresponded to a speed maximum. Because of these different patterns it is a nontrivial problem on how to define the middle of the vowel articulation kinematically. Further, the same kinds of problems occur when defining the middle of the vowel acoustically, because there is not necessarily an apparent local maximum or minimum in the formant trajectories sometime during vowel production. For the 1–2 secant line, the two kinematic quantities that were the least often close to a release or closure times were times of the maximum midpoint-to-palate distance and times of tangential acceleration zero crossing. (In cases where there were multiple zero crossings in the tangential acceleration, the one closest to the acoustic temporal center of the vowel was chosen.) Figures 3(a) and 3(b) show the proportion of total vowel time where these critical times for the 1–2 secant line occur as a function of vowel duration for both /i/ and /ε/ for each subject. The times of zero crossings of tangential accelerations were more consistently near the acoustic temporal centers than were the times of the maximum midpoint-to-palate distance for the 1–2 secant lines. Further, the 2–3 midpoint-to-palate distance maxima were not always well defined because this distance often varied little over the course of a vowel and because there were sometimes multiple local maxima. Thus, the zero crossings of the tangential acceleration were used for the kinematically defined midvowel, or vowel critical time. The phrase “vowel critical time” will sometimes be shortened to “vct” here.

#### B. Articulation of /i/ at vowel critical times and vowel duration

Figure 4(a) shows the positions of the 1–2 secant lines and the positions of the 2–3 secant lines at their respective vowel critical times from tokens of the word “this” for subject JW11. The midpoints of the secant lines are indicated as well. For JW11 there was nearly a continuum of 1–2 secant line midpoints at the 1–2 vowel critical times with continuous variation in their distances to the palate, except for one outlier. Also, the 2–3 secant line midpoints clustered together, except for one outlier. The outliers for both the 1–2 and 2–3 secant lines were from the same token in sentence

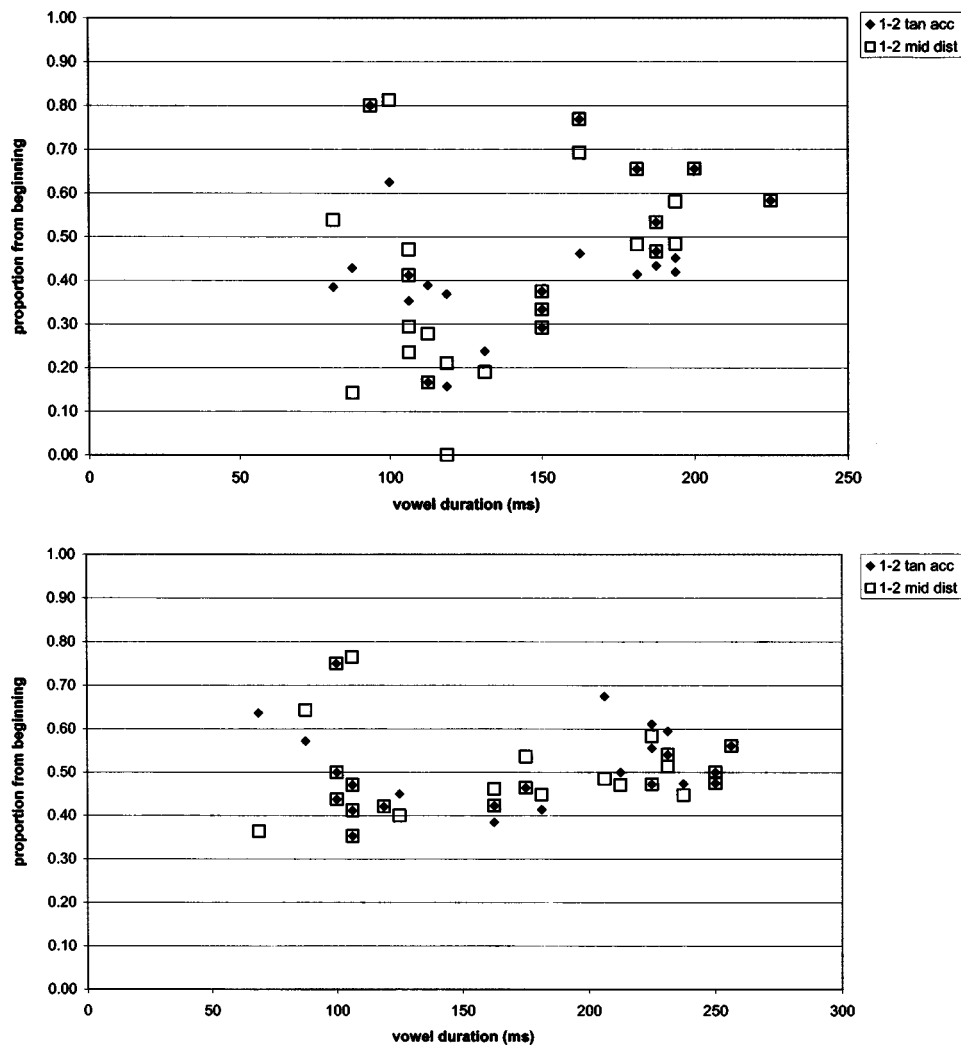


FIG. 3. Proportion of the total vowel duration where the vowel critical time occurred versus the vowel duration for all tokens of /ɪ/ and /ɛ/: (a) JW11; (b) JW29.

context: “...damage this ship’s hull.” It appears that this /ɪ/ was highly coarticulated with the /dʒ/ of the preceding word. The other /ɪ/ in sentence context had a 1–2 secant line midpoint that was the closest to the palate in the continuum seen in Fig. 4(a).

For JW29 there were clearly two groups of 1–2 secant lines at their vowel critical times [Fig. 4(b)]. These two groups were, perhaps, less clearly defined for the 2–3 secant midpoints. The tokens of the group of four 1–2 secant lines closest to the palate were from the word “this” spoken in sentence context, with vowel durations from 88 to 106 ms (Table I). The five 1–2 secant lines farthest from the palate were from the word “this” spoken in word lists, with vowel durations from 225 to 256 ms (Table I).

The first column of Table II contains the linear regression statistics for the secant line midpoint-to-palate distance at the vowel critical times versus vowel duration and for secant line orientation at the vowel critical times versus vowel duration for both subjects. The same statistics for the differences of these quantities with their values at release and closure are shown in the second and third columns of this table. Both JW11 and JW29 exhibited, at least, very significant ( $p < 0.05$ ) positive slopes in the linear regression of the 1–2 secant line midpoint-to-palate distance at the 1–2 vcts versus the vowel duration (Table II, rows 1 and 2, column 1),

i.e., the longer the duration the greater the distance between the palate and the midpoint of the secant lines. A highly significant positive slope occurred for JW29’s 2–3 secant line’s midpoint-to-palate distance versus vowel duration (Table II, row 4, column 1), but there was no such trend for JW11 (Table II, row 3, column 1). The differences in secant line midpoint-to-palate distances at vct and at either release or closure are shown in rows 1–4, columns 2 and 3 of Table II. Comparing across rows 1–4, and with the exception of JW11’s 2–3 secant line in row 3, the regression slopes in column 2 were less significant than those of column 1, and the regression slopes of column 3 were more significant than those in column 1. Thus, generally, differences in the midpoint-to-palate at vct and at closure followed the expected trend with vowel duration with more statistical significance than the midpoint-to-palate distance at either vct or its difference at vct and at release.

As noted in Sec. III A, the negative rotation of the tongue blade into a more downward pointing direction was a commonly observed pattern. It might be expected that the longer vowels would involve more of this negative orientation change than the shorter vowels. For JW11’s 1–2 secant line, it was only when the differences in orientation at vowel critical times and at release or closure that significant negative linear trends with vowel duration appeared (Table II,

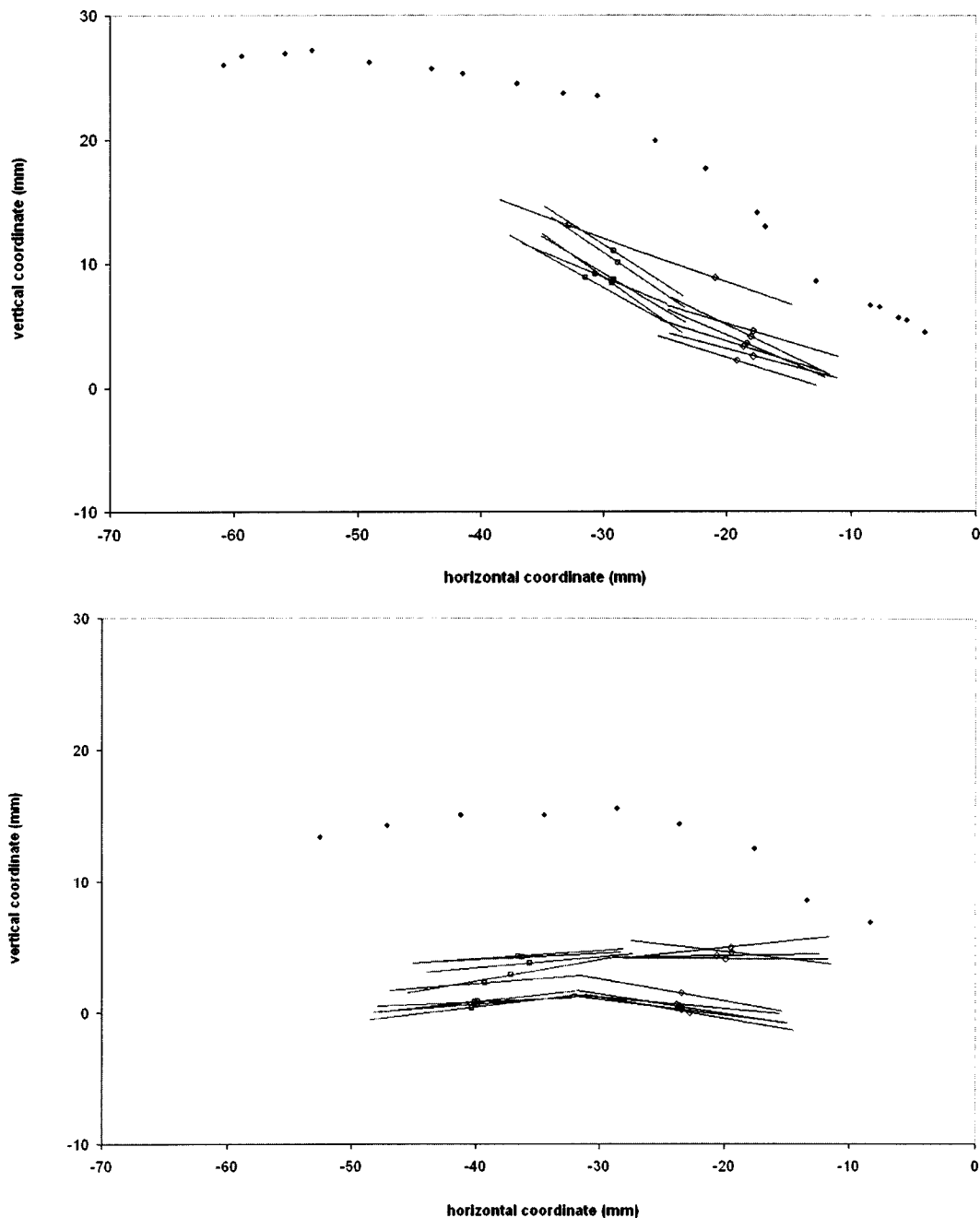


FIG. 4. 1–2 secant lines at their vowel critical times and 2–3 secant lines at their vowel critical times for tokens of “this.” The midpoints of the 1–2 secant lines are drawn as diamonds and the midpoints of the 2–3 secant lines are drawn as squares. (a) JW11, (b) JW29.

row 5). It was JW11’s 2–3 secant line orientation at the 2–3 vowel critical times or its difference with its value at release or closure that showed very significant negative trends with vowel duration (Table II, row 7). It can be noted that removing the outlier for JW11 did not significantly increase the degree of significance of any of these regression slopes. For JW29 the 1–2 secant line orientations at vcts, as well as differences of 1–2 orientations between vowel critical times and release or closure times, exhibited either very significant or highly significant negative slope with vowel duration (Table II, row 6). JW29’s 2–3 secant line orientation versus vowel duration showed significance only when differences between values at vct and closure were considered (Table II, row 8). Also, this trend was a *positive* slope. Other than for

JW29’s 2–3 secant line, the trend in orientation was for a more downward orientation with increased vowel duration.

### C. Formant frequencies and their relation to articulation of /ɪ/

Acoustic vowel reduction of /ɪ/ was not as evident for JW11, as it was for JW29. According to linear regression statistics, JW11 exhibited significance ( $p < 0.1$ ) only in a linear increase of F1 at 1–2 vowel critical times with vowel duration (slope=4.17,  $R^2 = 0.46$ ). Linear regression statistics for JW29 showed highly significant linear increases in F1 at vcts with increased vowel duration (slope=9.73,  $R^2 = 0.89$  at

TABLE II. Statistics for linear regression of the midpoint-to-palate distance and orientation at vowel critical time and the differences of these quantities at vowel critical time and at release and at closure versus the duration of the vowel /t/ in “this” for JW11 and JW29 (\*=0.1 confidence, \*\*=0.05 confidence, and \*\*\*=0.01 confidence). Distances are in mm, orientations in degrees, and durations in ms. “vct” stands for the vowel critical time.

Subject	Regression	@ vct	@ vct-@ rel	@ vct-@ clo
JW11	1-2 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.019 $R^2=0.71^{**}$	slope=0.010 $R^2=0.38$	slope=0.021 $R^2=0.80^{***}$
JW29	1-2 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.035 $R^2=0.96^{***}$	slope=0.037 $R^2=0.88^{***}$	slope=0.033 $R^2=0.83^{***}$
JW11	2-3 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.001 $R^2=0.00$	slope=-0.023 $R^2=0.63^{**}$	slope=-0.009 $R^2=0.15$
JW29	2-3 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.016 $R^2=0.77^{***}$	slope=-0.008 $R^2=0.29$	slope=0.020 $R^2=0.73^{***}$
JW11	1-2 orientation <i>versus</i> vowel duration	slope=-0.016 $R^2=0.04$	slope=-0.125 $R^2=0.56^*$	slope=-0.060 $R^2=0.64^{**}$
JW29	1-2 orientation <i>versus</i> vowel duration	slope=-0.053 $R^2=0.64^{***}$	slope=-0.089 $R^2=0.78^{***}$	slope=-0.044 $R^2=0.50^{**}$
JW11	2-3 orientation <i>versus</i> vowel duration	slope=-0.103 $R^2=0.82^{***}$	slope=-0.073 $R^2=0.69^{**}$	slope=-0.100 $R^2=0.91^{***}$
JW29	2-3 orientation <i>versus</i> vowel duration	slope=-0.004 $R^2=0.02$	slope=-0.001 $R^2=0.00$	slope=0.028 $R^2=0.46^{**}$

1-2 vct and slope=10.36,  $R^2=0.86$  at 2-3 vct), while she did not show any significant linear trend for F2 versus vowel duration.

One of the main reasons for using secant lines is that they can provide some insight into the articulatory acoustic relations. Table III provides statistics on linear regressions between F1 or F2 and the midpoint-to-palate distance or secant line orientation at vowel critical times for both subjects. For JW29, F1 significantly increased with both 1-2 secant line midpoint-to-palate distances at 1-2 vct and 2-3 secant line midpoint-to-palate distances at 2-3 vct (Table III, column 3, rows 1 and 2). There were no other significant linear

trends, except between F2 and the orientation of the 2-3 secant line for JW11.

#### D. Articulation of /ε/ at vowel critical times and vowel duration

In considering the vowel /ε/, the tokens of “sense” and “said” were combined for each speaker. 1-2 secant line midpoint-to-palate distances at vowel critical times, at release, and at closure are plotted as a function of the vowel duration of JW11’s and JW29’s /ε/ productions in Figs. 5(a) and 5(b). The figures indicate increased 1-2 secant line

TABLE III. Statistics for the linear regression of the secant line midpoint-to-palate distance and orientation versus formant frequencies at the vowel critical times (vct) for JW11 and JW29 producing /t/ in “this” (\*=0.1 confidence, \*\*=0.05 confidence, and \*\*\*=0.01 confidence). Distances are in millimeters, orientations in degrees, and frequencies in Hz.

<i>Versus</i>	JW11 F1	JW11 F2	JW29 F1	JW29 F2
1-2 midpoint-to-palate distance	slope=0.0011 $R^2=0.27$	slope=0.0010 $R^2=0.28$	slope=0.0020 $R^2=0.86^{***}$	slope=-0.0006 $R^2=0.03$
2-3 midpoint-to-palate distance	slope=0.0015 $R^2=0.34$	slope=-0.0009 $R^2=0.35$	slope=0.0009 $R^2=0.72^{***}$	slope=-0.0009 $R^2=0.16$
1-2 orientation	slope=-0.04 $R^2=0.09$	slope=-0.03 $R^2=0.17$	slope=-0.03 $R^2=0.61$	slope=-0.01 $R^2=0.05$
2-3 orientation	slope=-0.04 $R^2=0.09$	slope=-0.07 $R^2=0.69^{**}$	slope=-0.02 $R^2=0.02$	slope=-0.01 $R^2=0.07$

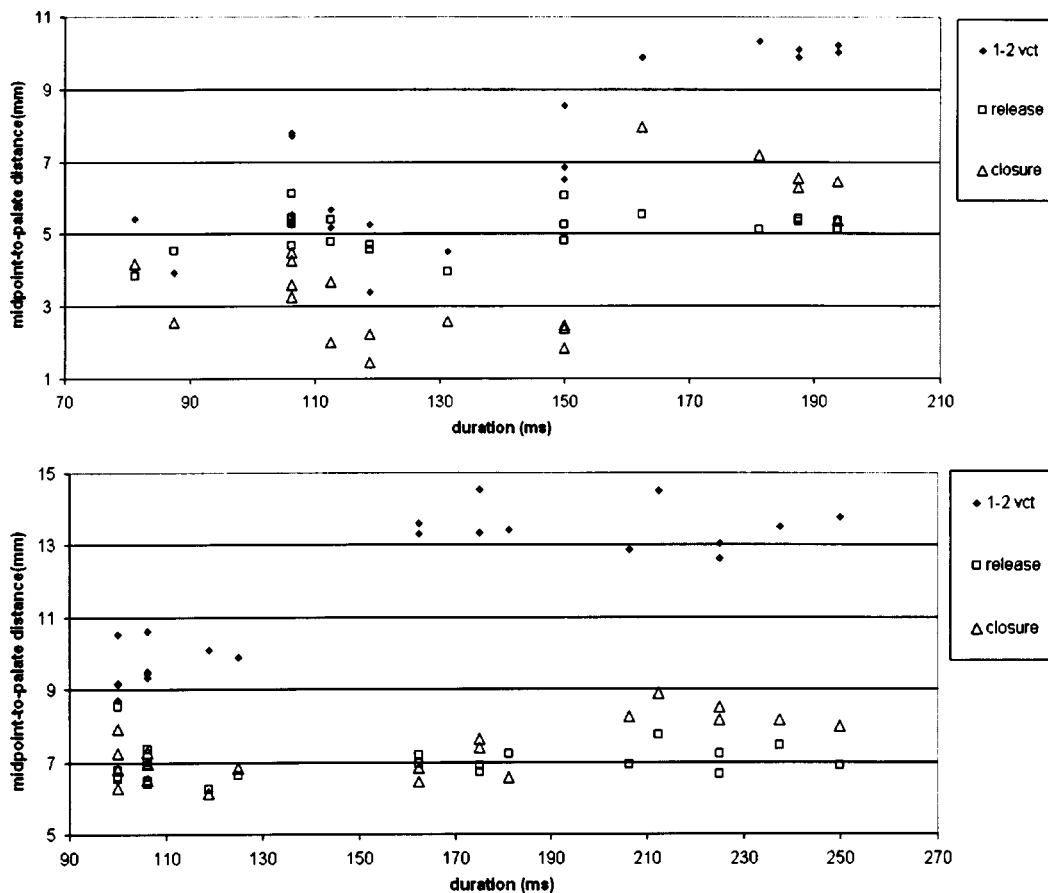


FIG. 5. The 1–2 midpoint-to-palate distance (mm) at release, at closure, and at vowel critical time versus vowel duration (ms) for tokens of /ε/ in the words “sense” and “said”. (a) JW11, (b) JW29.

midpoint-to-palate distance at vct with increased vowel duration for both subjects. However, variability in these distances for JW11 at release and at closure makes it difficult to find a trend visually in the differences at vct and at release and closure for this subject [Fig. 5(a)]. For both JW11 and JW29 the slopes of the regression lines for the 1–2 midpoint-to-palate distances at vcts versus vowel duration were all highly significant in the expected positive direction (Table IV, rows 1 and 2, column 1). Similarly, the linear trends between differences in the midpoint-to-palate distance measures taken at vowel critical times and at release or closure and vowel duration were, at least, very significant (Table IV, rows 1 and 2, columns 2 and 3). Also, the midpoint-to-palate distance at vct increased with vowel duration for JW11’s 2–3 secant line, as did the differences of this distance between vct and release or closure (Table IV, row 3). JW29’s 2–3 secant line behaved differently, with an increase in the midpoint-to-palate distance between vct and release with vowel duration, and a decrease when differences of these distances at vct and at closure were considered (Table IV, row 4, columns 1 and 3).

Secant line orientations at vowel critical times, at release, and at closure for the 1–2 secant lines of JW11’s and JW29’s /ε/ productions are shown in Figs. 6(a) and 6(b). A visual inspection of the figure does not reveal any particular trend for JW11, but there does seem to be a trend toward decreasing orientation with vowel duration at for JW29 [Fig. 6(b)]. For JW11 [Fig. 6(a)] there was some statistical indica-

tion of decreasing orientation at vct and more negative differences in orientation at vct and at release and closure (Table IV, row 5). These trends were statistically more significant for JW29 than for JW11 (Table IV, rows 5 and 6). As in the case of JW11’s /i/ utterances, there were very significant linear trends 2–3 secant line orientation at vct and differences with this quantity at release versus vowel duration (Table IV, row 7, columns 1 and 2). The orientation of JW29’s 2–3 secant line orientations at vowel critical times *increased* as a function of vowel duration, as did differences with orientations at release (Table IV, row 8, columns 1 and 2).

A closer look at Figs. 5(b) and 6(b) revealed that for JW29’s all vowels with durations greater than about 150 ms behaved similarly in the 1–2 secant line midpoint-to-palate distance and orientation measures. This was a saturation effect that means that a linear fit to the data did not capture all the regularity in the 1–2 secant line distance to palate and orientation versus vowel duration relations. A more appropriate function fit might be provided by an exponential function of vowel duration; similar to the one that Lindblom (1963) used to model behavior of the formant frequencies in mid-vowel as a function of vowel duration.

### E. Formant frequencies and their relation to articulation of /ε/

There were individual differences in the way formant frequencies at vowel critical times behaved as a function of

TABLE IV. Statistics for linear regression of the midpoint-to-palate distance and orientation at vowel critical time and the differences of these quantities at vowel critical time and at release and at closure versus the duration of the vowel /ε/ in “sense and said” for JW11 and JW29 (\*=0.1 confidence, \*\*=0.05 confidence, and \*\*\*=0.01 confidence). Distances are in mm, orientations in degrees, and durations in ms. “vct” stands for vowel critical time.

Subject	Regression	@ vct	@ vct—@ rel	@ vct—@ clo
JW11	1–2 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.053 $R^2=0.69^{***}$	slope=0.047 $R^2=0.72^{***}$	slope=0.020 $R^2=0.32^{**}$
JW29	1–2 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.033 $R^2=0.74^{***}$	slope=0.032 $R^2=0.39^{***}$	slope=0.024 $R^2=0.23^{**}$
JW11	2–3 midpoint-to-palate distance <i>versus</i> vowel duration	slope=0.042 $R^2=0.67^{***}$	slope=0.024 $R^2=0.31^{**}$	slope=0.025 $R^2=0.75^{***}$
JW29	2–3 midpoint-to-palate distance <i>versus</i> vowel duration	slope=-0.005 $R^2=0.06$	slope=0.087 $R^2=0.21^{**}$	slope=-0.013 $R^2=0.38^{***}$
JW11	1–2 orientation <i>versus</i> vowel duration	slope=-0.086 $R^2=0.16^*$	slope=-0.082 $R^2=0.20^{**}$	slope=-0.121 $R^2=0.26^{**}$
JW29	1–2 orientation <i>versus</i> vowel duration	slope=-0.144 $R^2=0.72^{***}$	slope=-0.159 $R^2=0.81^{***}$	slope=-0.119 $R^2=0.65^{***}$
JW11	2–3 orientation <i>versus</i> vowel duration	slope=-0.126 $R^2=0.40^{***}$	slope=-0.206 $R^2=0.56^{***}$	slope=0.028 $R^2=0.04$
JW29	2–3 orientation <i>versus</i> vowel duration	slope=0.071 $R^2=0.77^{***}$	slope=0.027 $R^2=0.15^*$	slope=-0.001 $R^2=0.03$

vowel duration. For JW29, F1 showed highly significant linear trends with vowel duration at both the 1–2 and 2–3 secant line vowel critical times, just as for her /i/ tokens (slope=11.52,  $R^2=0.82$ ,  $p<0.01$  at 1–2 vct, and slope=10.39,  $R^2=0.72$ ,  $p<0.01$  at 2–3 vct). For JW11, the linear relation between F1 and vowel duration was significant only at the 2–3 secant line vcts (slope=5.24,  $R^2=0.26$ ,  $p<0.05$  at 2–3 vct). However, only JW11 showed a highly significant linear trend in F2 with vowel duration in /ε/ (slope=6.78,  $R^2=0.34$ ,  $p<0.01$  at 1–2 vct, and slope=11.29,  $R^2=0.32$ ,  $p<0.01$  at 2–3 vct).

Further insight is gained when the dependencies of the formant frequencies on the articulatory measures are examined. F1 of JW11 showed a significant dependence on the 1–2 secant line and a very significant dependence on the 2–3 secant line midpoint-to-palate distances (Table V, rows 1 and 2, column 1). The linear trend of JW11’s F1 versus secant line orientation was negative and highly significant for both the 1–2 and 2–3 secant lines (Table V, rows 3 and 4, column 1). The F2 versus midpoint-to-palate distance and F2 versus orientation relations were, at least, very significant for each of JW11’s secant lines (Table V, column 2). For JW29’s F1

there was a highly significant dependence on both the 1–2 secant line midpoint-to-palate distance and orientation (Table V, rows 1 and 3, column 3). There was little dependence of JW29’s F2 on any articulatory quantity, except for a significant negative slope for F2 versus 1–2 orientation. A significant *negative* trend in F1 versus the 2–3 secant line midpoint-to-palate distance was recorded (Table V, row 2, column 3).

There is further evidence of individual differences in articulatory-acoustic relations in /ε/ production when horizontal and vertical dimensions of the midpoints are considered separately. The secant line that shows the most significant trend in the midpoint-to-palate distance versus F1 is the 2–3 secant line for JW11 and the 1–2 secant line for JW29. For JW11, a highly significant linear relation exists between the horizontal coordinate of the 2–3 secant line midpoint and F1 (slope=-0.016,  $R^2=0.48$ , and  $p<0.01$ ), but not for the vertical coordinate at the 2–3 secant line vct (slope=-0.010,  $R^2=0.13$ ). For JW29, the linear fits between F1 and the horizontal coordinate and between F1 and the vertical coordinate were both highly significant (slope=-0.013,  $R^2=0.65$ , and  $p<0.01$  and slope=-0.010,  $R^2=0.36$ , and  $p<0.01$ , respectively).

#### IV. DISCUSSION

The results indicate that phonetic vowel reduction occurred for both subjects, to varying degrees, as measured in terms of midpoint-to-palate distances, orientations, and formant frequencies. After reviewing the important, particular facts about the dataset examined here, we will summarize the findings regarding articulatory changes with vowel duration. The second part of the discussion will focus on the physics of the observed relations between articulatory and acoustic trends. This will entail some discussion on individual differences.

There are three facts that have relevance to the following discussion. First, the limited range of some of the observations could have had an effect on the linear regressions. In some cases a lack of sufficient range in the midpoint-to-palate or secant line orientation could have made it difficult to find significant linear trends, because the variability due to noise might have outweighed variability that resulted in regular linear trends. The lack of range could have been the result of a compensatory articulatory strategy, as discussed for JW11’s /i/ in Sec. III A. Second is the fact that the same secant line probably spanned different functional regions of the tongue for the two subjects, as discussed in Sec. II. The final fact is that the subjects had substantially different palate shapes.

The following general trends in articulation as a function of vowel duration have been noted using Tables II and IV. The 1–2 secant line midpoint-to-palate distances at vct increase with vowel duration for /i/ and /ε/ for both subjects. For JW11’s 2–3 secant line there was no such trend for /i/, but there was for /ε/. The lack of a trend for JW11’s /i/ could have been due to his limited range of vowel durations for this vowel. For JW29’s 2–3 secant line, there was no consistent trend in the midpoint-to-palate distance for /ε/, but there was

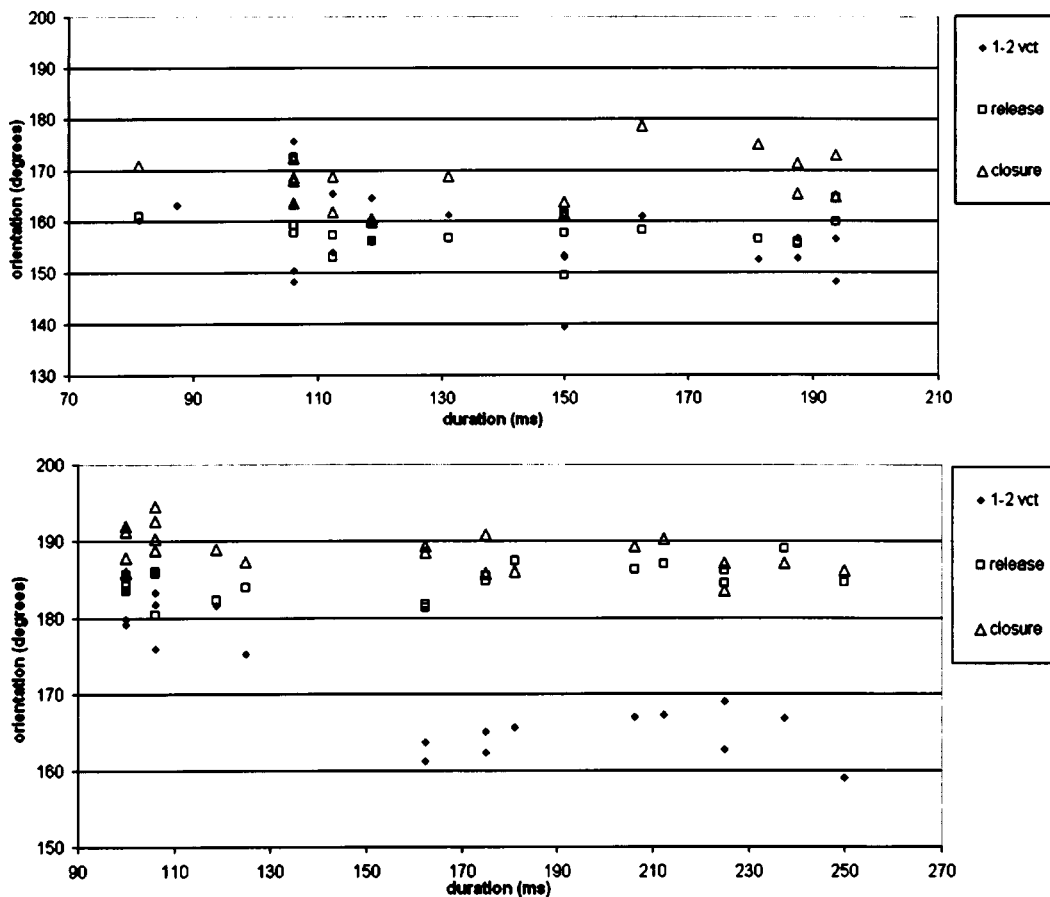


FIG. 6. 1–2 secant line orientation (degrees) at release, at closure, and at 1–2 vowel critical time versus vowel duration (ms) for tokens of /ε/ in the words “sense” and “said”. (a) JW11; (b) JW29.

for /ɪ/. The differences between the subjects in the behavior of the 2–3 secant line was highlighted when orientations were considered.

JW29’s 1–2 secant line showed increased downward (i.e., more negative) orientation with increased vowel duration to higher degrees of significance than did JW11’s 1–2 secant line for /ɪ/ and /ε/. The opposite is true for the 2–3 secant line, with JW11’s showing a more significant increase in downward orientation with vowel duration than JW29’s, for both vowels. In fact, JW29’s 2–3 secant line had some significant increases in orientation with vowel duration. The differences in the behavior of JW29’s and JW11’s 2–3 secant

line could be attributed to a difference in the functional parts of the tongue spanned by this secant line for the two subjects. Palate shape differences could also have played a role.

We now turn to the relation between articulation and acoustics reported in Tables III and V. There were no significant trends between F1 and the midpoint-to-palate distance for JW11 in the case of /ɪ/, while there were highly significant trends for JW29 (Table III). This could have been the result of his limited range of midpoint-to-palate distances for this subject. The ranges of midpoint-to-palate distances for JW11’s /ɪ/ productions were less than 3.5 mm for both the 1–2 and 2–3 secant lines, while the range for JW29’s 1–2

TABLE V. Statistics for linear regression of the secant line midpoint-to-palate distance and orientation versus formant frequencies at the vowel critical times (vct) for JW11 and JW29 producing /ε/ in “sense and said” (\*=0.1 confidence, \*\*=0.05 confidence, and \*\*\*=0.01 confidence). Distances are in millimeters, orientations in degrees, and frequencies in Hz.

Versus	JW11 F1	JW11 F2	JW29 F1	JW29 F2
1–2 midpoint-to-palate distance	slope= 0.0021 $R^2=0.16^*$	slope= 0.0024 $R^2=0.50^{***}$	slope= 0.0015 $R^2=0.60^{***}$	slope= 0.0013 $R^2=0.09$
2–3 midpoint-to-palate distance	slope= 0.0017 $R^2=0.28^{**}$	slope= 0.0009 $R^2=0.32^{***}$	slope= 0.0004 $R^2=0.15^*$	slope= 0.0000 $R^2=0.00$
1–2 orientation	slope=-0.11 $R^2=0.41^{***}$	slope=-0.06 $R^2=0.24^{**}$	slope=-0.06 $R^2=0.54^{***}$	slope=-0.07 $R^2=0.15^*$
2–3 orientation	slope=-0.08 $R^2=0.45^{***}$	slope=-0.05 $R^2=0.70^{***}$	slope=-0.02 $R^2=0.22^{**}$	slope=0.01 $R^2=0.02$

secant line was 6.9 and 3.6 mm for her 2–3 secant line in /i/ (see Figs. 4 and 5). In the case of JW29's /i/ productions, these trends in F1 versus midpoint-to-palate distances provide examples of the well-known relation between tongue height and F1. An increase in the area of the front part of the vocal tract can be conceived as a decrease in the magnitude of the mass element of the Helmholtz resonator for high articulations. For less close articulations, increases in the midpoint-to-palate distance represented increases in the portion of the acoustic tube that contains the volume velocity maximum for the standing wave associated with F1, in an application of perturbation theory (e.g., Schroeder, 1967; Stevens, 1998, pp. 148–152).

For /ε/ the positive F1 versus midpoint-to-palate distance trend was most significant for the 2–3 secant line of JW11 and the 1–2 secant line of JW29. The difference between subjects appears partly to be due to the differences in functional regions spanned by their 2–3 secant lines. However, it can be noted that changes in F1 were not strictly related to tongue height, because the shape of the palate plays a role. The fact that the relation between the horizontal coordinate of JW11's 2–3 secant line midpoint and F1 was highly significant, but there was no significant relation between the vertical component and F1, could be due to the fact that the 2–3 secant line opposed the relatively steep diagonal of JW11's palate. Thus, pulling the tongue back for JW11's /ε/ production had the same effect as JW29's pulling the tongue down and back from her flatter palate: an increase in F1.

The palate shape differences could also explain why there were significant increases of F2 with midpoint-to-palate distances for JW11 in the vowel /ε/ and no significant trends between midpoint-to-palate distances and F2 for JW29 in either vowel. The highly significant increase of JW11's F2 in /ε/ with 1–2 and 2–3 midpoint-to-palate distances could have been caused in two ways. The 1–2 and 2–3 segments of JW11 could have been far enough forward that their translations affected only the maximum velocity portion of the F2 standing wave, or the increased midpoint-to-palate distance of these segments was coincident with a decreasing orientation angle, which actually caused the F2 increase. The reason to expect an increased F2 with a decreased orientation angle is explained below.

Based on perturbation theory, it can be expected that both F1 and F2 of /ε/ increase with decreasing orientation angle for secant lines in the front of the tongue. Decreasing orientation angles in the front secant lines imply that the regions for velocity maxima of the F1 and F2 have an increasing cross-sectional area, while the regions posterior to this, which could be associated with either regions of pressure maxima or regions that are neither velocity or pressure maxima, possess decreasing areas. There were the expected, and very significant, negative trends between F1 and 1–2 and 2–3 secant line orientation in the vowel /ε/ for JW11 and JW29, and in the case of JW11, between F2 and 1–2 and 2–3 secant line orientation (Table V). However, it must be kept in mind that these effects could simply be due to the coincidence of increasing the midpoint-to-palate distance and decreasing orientation, with the physical cause for

changing F1 and F2 being the increasing midpoint-to-palate distance. The significances of trends for /i/ (Table III) were probably hindered by a lack of range of orientations. For JW11, the ranges of orientations for both secant lines in /i/ were less than 15°, compared to ranges greater than 25° for /ε/. The JW29 ranges of orientations for both secant lines in /i/ were less than 16°, compared to ranges greater than 17° for /ε/. The fact that the 2–3 secant line appeared to have consistent effects on F1 and F2 for JW11, but not for JW29, may have been the result of the fact that the functional extent of the tongue blade accounted for between the first and third pellets is greater for JW29 than for JW11.

In summary, while we have claimed that pellet placement did have an effect on how tongue placement in vowel production was measured, we believe that there were speaker differences that cannot be attributed to any differences to pellet placement differences. In effect we have added a small amount of evidence for the conclusion of Johnson, Ladefoged, and Lindau, "*We are left with the conclusion that speakers of the same language and dialect may use different articulatory plans; that the universal articulatory phonetics hypothesis is wrong*" (Johnson *et al.*, 1993, p. 713). It can be seen here that some of this difference could be the result of palate shape differences. This also has implications for the normalization procedure of Hashi, Westbury, and Lindstrom (1998), which essentially maps the midsagittal palate outlines onto a straight line for all speakers. They found that the variation in the pellet-to-palate distance between talkers was reduced with this mapping. We have an analogous normalization with the secant line midpoint-to-palate distance measure. Our results indicate that if such a normalization procedure is to be employed, then secant line orientations need to be measured in relation to the palate, and the oral space near the palate needs to be warped to account for directions of translational tongue movement in relation to the local palate surface.

## V. CONCLUSION

In this specialized sample of lax, front vowels in alveolar or dental consonant context, both midpoint-to-palate distance differences and orientation differences contributed to acoustic assimilation to the surrounding consonants with decreasing vowel duration. Further, speaker-dependent differences in articulation and the formant frequency-articulation relations appeared to be, in part, attributable to palate shape differences. Secant line coordinates is a representation that is closely related to flesh point data that successfully emphasizes the relation between local tongue shape and acoustics.

There are at least two directions for current research to continue. The first is to examine the same phenomena in other, so-called, front vowels, such as /i/ and /æ/, and to treat linguistic factors, such as stress, separately from duration. Other extensions involve studying the properties of the secant line trajectories themselves. Data-driven articulatory synthesis should be achievable through the control of secant line kinematics.

## ACKNOWLEDGMENTS

This work was supported by Grant No. NIDCD-01247 to CReSS LLC. The comments of Anders Löfqvist, Tine Mooshammer, and an anonymous reviewer helped to greatly improve this paper.

- Adams, S. G., Weismer, G., and Kent, R. D. (1993). "Speaking rate and speech velocity movement profiles," *J. Speech Hear. Res.* **36**, 41–54.
- van Bergem, D. R. (1993). "Acoustic vowel reduction as a function of sentence accent, word stress, and word class," *Speech Commun.* **12**, 1–23.
- Broad, D. J., and Fertig, R. H. (1970). "Formant-frequency trajectories in selected CVC-syllable nuclei," *J. Acoust. Soc. Am.* **47**, 1572–1582.
- Broad, D. J., and Clermont, F. (1987). "A methodology for modeling vowel formant contours in CVC context," *J. Acoust. Soc. Am.* **81**, 155–165.
- Di Benedetto, M.-G. (1989). "Vowel representation: Some observations on temporal and spectral properties of the first formant frequency," *J. Acoust. Soc. Am.* **86**, 55–66.
- Engstrand, O. (1988). "Articulatory correlates of stress and speaking rate in Swedish VCV utterances," *J. Acoust. Soc. Am.* **83**, 1863–1875.
- Flège, J. E. (1988). "Effect of speaking rate on tongue position and velocity of movement," *J. Acoust. Soc. Am.* **84**, 901–916.
- Fujimura, O., Kiritani, S., and Ishida, H. (1973). "Computer controlled radiography for observation of movements of articulatory and other human organs," *Comput. Biol. Med.* **3**, 371–384.
- Gay, T. (1978). "The effect of speaking rate on vowel formant movements," *J. Acoust. Soc. Am.* **63**, 223–230.
- Harris, K. S. (1978). "Vowel duration change and its underlying physiological mechanisms," *Lang Speech* **21**, 354–361.
- Hashi, M., Westbury, J. R., and Honda, K. (1998). "Vowel posture normalization," *J. Acoust. Soc. Am.* **104**, 2426–2437.
- Hoole, P. (1999). "On the lingual organization of the German vowel system," *J. Acoust. Soc. Am.* **106**, 1020–1032.
- Johnson, K., Ladefoged, P., and Lindau, M. (1993). "Individual differences in vowel production," *J. Acoust. Soc. Am.* **94**, 701–714.
- Keating, P. A., Lindblom, B., Lubker, J., Kreiman, J. (1994). "Variability in jaw height for segments in English and Swedish VCVs," *J. Phonetics* **22**, 407–422.
- Kuehn, D. P., and Moll, K. L. (1976). "A cineradiographic study of VC and CV articulatory velocities," *J. Phonetics* **4**, 303–320.
- Lindblom, B. (1963). "Spectrographic study of vowel reduction," *J. Acoust. Soc. Am.* **35**, 1773–1781.
- Lindblom, B. (1964). "Articulatory activity in vowels," *KTH Speech Technology Laboratory-QPSR*, Vol. 2/1964, pp. 1–5.
- Milenkovic, P. (2003). Personal communication.
- Moon, S.-J., and Lindblom, B. (1994). "Interaction between duration, context, and speaking style in English stressed vowels," *J. Acoust. Soc. Am.* **96**, 40–55.
- Nord, L. (1986). "Acoustic studies of vowel reduction in Swedish," *KTH Speech Transmission Laboratory-QPSR*, Vol. 4/1989, pp. 19–36.
- Ohde, R. N., and Scharf, D. J. (1975). "Coarticulatory effects of voiced stops on the reduction of acoustic targets," *J. Acoust. Soc. Am.* **58**, 923–927.
- Ostry, D. J., and Munhall, K. G. (1985). "Control of rate and duration of speech movements," *J. Acoust. Soc. Am.* **77**, 640–648.
- Perkell, J. S. (1991). "On the nature of distinctive features: Implications of a preliminary vowel production study," in *Frontier of Speech Communication Research*, edited by B. Lindblom and S. Öhman (Academic, London), pp. 365–380.
- Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., and Jackson, M. T. T. (1992). "Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements," *J. Acoust. Soc. Am.* **92**, 3078–3096.
- Perkell, J. S., Zandipour, M., Matthies, M. L., and Lane, H. (2002). "Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues," *J. Acoust. Soc. Am.* **112**, 1627–1641.
- Schroeder, M. R. (1967). "Determination of the geometry of the human vocal tract by acoustic measurements," *J. Acoust. Soc. Am.* **41**, 1002–1010.
- Simpson, A. P. (2002). "Gender-specific articulatory-acoustic relations in vowel sequences," *J. Phonetics* **30**, 417–435.
- van Son, R. J. J. H., and Pols, L. C. W. (1990). "Formant frequencies of Dutch vowels in a text read at normal and fast rate," *J. Acoust. Soc. Am.* **88**, 1683–1693.
- van Son, R. J. J. H., and Pols, L. C. W. (1992). "Formant movements of Dutch vowels in a text read at normal and fast rate," *J. Acoust. Soc. Am.* **92**, 121–127.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
- Stevens, K. N., and House, A. S. (1963). "Perturbation of vowel articulation by consonantal context: An acoustical study," *J. Speech Hear. Res.* **6**, 11–128.
- Stevens, K. N., House, A. S., and Paul, A. P. (1966). "Acoustic description of syllable nuclei: An interpretation in terms of a dynamic model of articulation," *J. Acoust. Soc. Am.* **40**, 123–132.
- Stone, M. (1990). "Three-dimensional tongue movement," *J. Acoust. Soc. Am.* **87**, 2207–2217.
- Westbury, J. R. (1994). *X-Ray Microbeam Speech Production Database User's Handbook* (unpublished manuscript, University of Wisconsin).
- Westbury, J. R., Hashi, M., and Lindstrom, M. J. (1998). "Differences among speakers in lingual articulation for American English /ɪ/," *Speech Commun.* **26**, 203–226.