

Subglottal coupling and its influence on vowel formants

Xuemin Chi^{a)} and Morgan Sonderegger^{b)}

Speech Communication Group, RLE, MIT, Cambridge, Massachusetts 02139

(Received 25 September 2006; revised 14 June 2007; accepted 15 June 2007)

A model of acoustic coupling between the oral and subglottal cavities is developed and predicts attenuation of and discontinuities in vowel formant prominence near resonances of the subglottal system. One discontinuity occurs near the second subglottal resonance (*SubF2*), at 1300–1600 Hz, suggesting the hypothesis that this is a quantal effect [K. N. Stevens, *J. Phonetics* **17**, 3–46 (1989)] dividing speakers' front and back vowels. Recordings of English vowels (in /hVd/ environments) for three male and three female speakers were made, while an accelerometer attached to the neck area was used to capture the subglottal waveform. Average speaker *SubF2* values range from 1280 to 1620 Hz, in agreement with prior work. Attenuation of 5–12 dB of second formant prominence near *SubF2* is found to occur in all back-front diphthongs analyzed, while discontinuities in the range of 50–300 Hz often occur, in good agreement with the resonator model. These coupling effects are found to be generally stronger for open-phase than for closed-phase measurements. The implications for a quantal relation between coupling effects near *SubF2* and [back] are discussed. © 2007 Acoustical Society of America. [DOI: 10.1121/1.2756793]

PACS number(s): 43.70.Bk [BHS]

Pages: 1735–1745

I. INTRODUCTION

An important question lies at the intersection of phonetics and feature-based phonology: What is the precise connection between the acoustic properties of sounds and their phonological feature values? One proposal addressing this question is quantal theory (QT) (Stevens, 1972, 1989), in which features arise from nonlinear articulatory-acoustic relations in speech production when an acoustic parameter of the speech signal shifts from one stable region to another for a linear articulatory movement (schematized in Fig. 1).

These are termed “quantal relations”; possible quantal relations for most distinctive features are given in Stevens (1989). Experimental studies that address QT (Stevens and Blumstein, 1975; Ladefoged and Bhaskararao, 1983; Pisoni, 1981; Perkell and Cohen, 1989) generally have not examined particular quantal relations, focusing instead on whether a phonetic data set is “quantal” in the broad sense that it shows variation in articulation of a particular sound structured so as to minimize variation in its acoustic realization. However, almost no work so far has directly tested the central claim of QT, that certain acoustic-articulatory nonlinearities are used as quantal relations. A direct test of QT would have to address two questions about a particular acoustic-articulatory relation proposed as a quantal relation for [feature]: First, does the nonlinearity occur robustly in real speech? Second, if it does, is it used by speakers to distinguish sounds with + and – values of [feature]? Without direct tests of individual quantal relations, it is unclear whether QT should be seen as a theoretical construct, a set of constraints on speech sounds, or the fundamental basis of phonetic categories. This paper takes a first step toward testing QT by analyzing a nonlinear

acoustic-articulatory relation proposed as a quantal relation in detail, and thus answering the first question, that of occurrence and robustness.

This paper examines second formant frequency discontinuity and amplitude attenuation in diphthongs near the second subglottal resonance (*SubF2*) due to coupling between the oral and subglottal cavities (“subglottal coupling”). Subglottal resonances have been measured in a number of studies using both invasive (Ishizaka *et al.*, 1976; Cranen and Boves, 1987) and noninvasive (Fant *et al.*, 1972; Henke, 1974; Stevens *et al.*, 1975; Hanson and Stevens, 1995; Cheyne, 2002) methods. *SubF2* usually lies in the range of 1300–1600 Hz, depending on the speaker. This frequency range is also approximately the dividing line between the second formant (*F2*) of front ([–back]) and back ([+back]) vowels. What determines the boundary, between front and back vowels is unknown. QT provides one possibility, that speakers will avoid putting formants in the acoustically unstable region around *SubF2*, which will then be used as a dividing line between front and back vowels.

In this paper, a theoretical model is first presented to explain the occurrence of subglottal coupling effects, followed by analysis of the occurrence and size of these effects, namely attenuation of the second formant amplitude (*A2*) and a discontinuity in second formant frequency near *SubF2*, and finally discussion of the implications for testing QT.¹

II. MODELING ORAL-SUBGLOTTAL COUPLING

Coupling between the oral and subglottal cavities can usually be ignored in acoustic modeling of the vocal tract transfer function, but becomes non-negligible when an oral cavity formant approaches a subglottal resonance in frequency. General properties of coupled resonators (discussed in the following) then predict that the formant prominence amplitude will be attenuated, and its prominence frequency will appear to “jump,” skipping the subglottal resonance fre-

^{a)}Electronic mail: xuemin@speech.mit.edu

^{b)}Electronic mail: morgan@speech.mit.edu

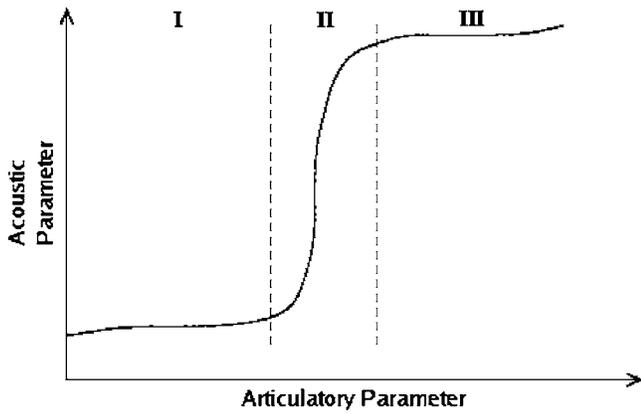


FIG. 1. Schematic of a quantal relation. Regions I and III correspond to [+feature] and [-feature] values, region II is the abrupt transition between them.

quency. The latter effect can often be observed as a discontinuity in the F_2 track for back-front diphthongs near the second subglottal resonance, as in Fig. 2. These attenuation and frequency jump effects, together called “subglottal coupling effects,” can be understood through the acoustic model of the oral-subglottal system described in the following. This model predicts formant frequency jumps and A_2 attenuation due to subglottal coupling, and suggests a positive correlation between the amount of a speaker’s subglottal coupling and the size of that speaker’s frequency jump and amplitude attenuation effects.

A. Description of the acoustic model

In this section, an acoustic model of coupling between the oral and subglottal cavities is developed in order to understand the occurrence of subglottal coupling effects seen in some back-front diphthongs. Interaction between the oral and subglottal cavities can be modeled by the circuit shown in Fig. 3.

In this circuit, the volume velocity output of the glottis, U_0 , is filtered by the circuit’s transfer function, $T(\omega)$, to give U_m , the volume velocity at the lips. This setup deviates from

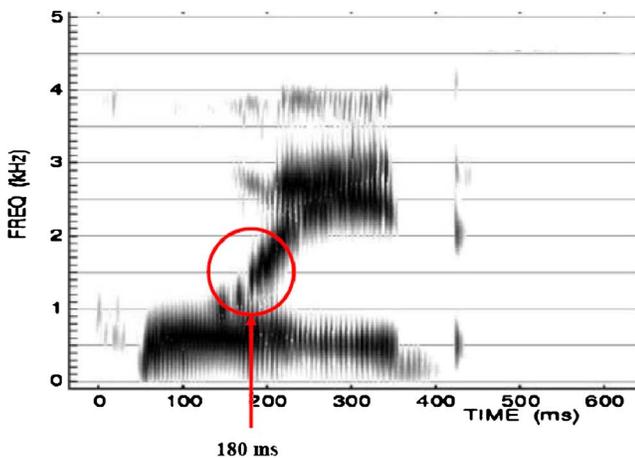


FIG. 2. (Color online) Spectrogram for an /aɪ/ diphthong spoken by a male speaker (M1). As shown in the circled region at 180 ms, attenuation of F_2 prominence and discontinuity in the F_2 track occur near the second subglottal resonance, at 1370 Hz.

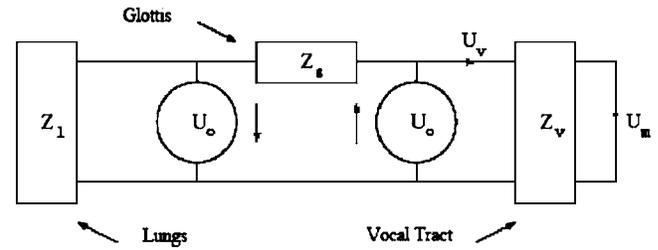


FIG. 3. Equivalent circuit model of the oral and subglottal cavities. Z denotes impedances. Adapted from Hanson and Stevens (1995).

the standard source-filter model only in incorporating the subglottal cavity, via coupling to the vocal tract through the glottis, into the transfer function, which is conventionally determined only by the vocal tract.

In Fig. 3, Z_l is the impedance of the subglottal system, approximated here by a tube of length l_l and area A_l terminated in a lossy capacitance (acoustic compliance) Z_c , following an approximation suggested in Ishizaka *et al.* (1976) to capture the damping effect of the lung bronchi. This approximation is acceptable when an accurate model of only one subglottal formant is required, as in this paper.²

Z_g is the glottal impedance, modeled by

$$Z_g = R_g + j\omega M_g, \quad (1)$$

with $R_g = 12\mu h_g / l_g d_g^3 + K_\rho U_g / (l_g d_g)^2$ and $M_g = \rho h_g / l_g d_g$, where μ and ρ are the glottal coefficients of viscosity and density, l_g, d_g, h_g are the glottal length, width, and thickness, U_g is the glottal volume velocity, and K is a constant taken to be ≈ 1 (Rösler and Strube, 1989; Stevens, 1998, p. 165). U_g is periodic, but since its period is small compared to the time scale of formant transitions, it can be approximated here by its average value without affecting the model behavior, allowing a linear system analysis. The dipole configuration of two volume velocity sources U_0 separated by impedance Z_g in Fig. 3 is a good approximation for the glottal source (Zhao *et al.*, 2002).

Z_v , the impedance looking into the vocal tract from the glottis, is computed here using the two-tube model of the vocal tract shown in Fig. 4. This model is sufficiently accurate to predict F_1 and F_2 , which characterize different vowels (along the height and backness dimensions). Vocal tract wall impedances are neglected for simplicity, but to give the vocal tract formants some bandwidth in the frequency range

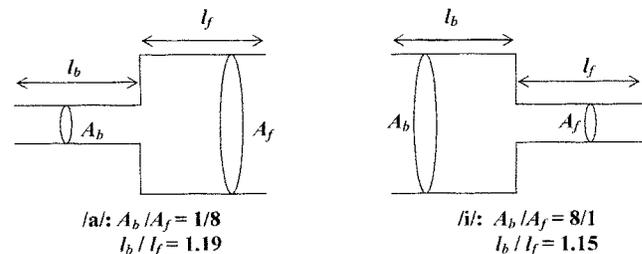


FIG. 4. Two-tube models of the vocal tract configuration for /a/ and /i/, adapted from Fant (1960), p. 66. For /a/, $A_b = 1 \text{ cm}^2$, $A_f = 8 \text{ cm}^2$, $l_b = 10.7 \text{ cm}$, and $l_f = 9.0 \text{ cm}$. For /i/, $A_b = 8 \text{ cm}^2$, $A_f = 1 \text{ cm}^2$, $l_b = 7.9 \text{ cm}$, and $l_f = 6.9 \text{ cm}$. Z_f is the impedance looking into the front cavity and Z_r is the radiation impedance at the lips.

of interest, an output resistance was added in series with the radiation impedance for the simulations described in Sec. II B.

In Fig. 4, Z_r is the radiation impedance at the lips and Z_f is the impedance looking into the front cavity. For the frequency range simulated here, Z_r can be approximated as

$$Z_r = \frac{\rho\omega^2}{4c} K_s(\omega) + j \frac{\omega\rho(0.8a)}{A_m}, \quad (2)$$

where $\pi a^2 = A_m$, the area of the mouth opening. $K_s(\omega)$ is a numerical factor accounting for the baffling effect of the head, which can be approximated as 1 at low frequencies, rising to 1.7 by 2000 Hz, and 1.5 for frequencies above 2000 Hz (Stevens, 1998, p. 153).

The impedance looking down a tube of length l and area A terminated in an impedance Z is (Kinsler and Frey, 1962, p. 201)

$$Z_0(l, A, Z) = \frac{\rho c}{A} \cdot \frac{Z + j \frac{\rho c}{A} \tan kl}{\frac{\rho c}{A} + j Z \tan kl}, \quad (3)$$

where c is the speed of sound. For the two tube model, in Fig. 4, Eq. (3) thus gives $Z_f = Z_0(l_f, A_f, Z_r)$ and $Z_v = Z_0(l_b, A_b, Z_f)$.

To determine the transfer function, $T(\omega) = U_m / U_0$, in Fig. 3, $T(\omega)$ is decomposed into

$$T(\omega) = \frac{U_m U_v}{U_v U_0}, \quad (4)$$

where U_v is the airflow into the vocal tract, and the two ratios of volume velocities are first solved individually.

U_m / U_v is determined by impedance matching at the entrances to the front and back cavities, which leads to

$$\frac{U_m}{U_v} = \frac{[K_b \cos(\omega l_b / c) - j \sin(\omega l_b / c) Z_f]}{K_b} \times \frac{[K_f \cos(\omega l_f / c) - j \sin(\omega l_f / c) Z_v]}{K_f}, \quad (5)$$

after some calculation, where $K_f = \rho c / A_f$ is the characteristic impedance of the front cavity and K_b is defined similarly.

Solving the circuit in Fig. 3 gives

$$\frac{U_v}{U_0} = \frac{Z_g}{Z_g + Z_v + Z_l}, \quad (6)$$

where U_v is the airflow into the vocal tract. Under the usual assumption of zero oral-subglottal coupling, $Z_g = \infty$, U_v / U_0 is unity and contributes no poles or zeros to the transfer function $T(\omega)$. As shown in Sec. II C, when coupling is assumed, numerical calculation of Eq. (6) using the expressions for Z_g , Z_v , and Z_l given earlier shows that a zero-pole pair is added to the transfer function at each subglottal resonance, with the zero at the subglottal resonance (when Z_l has a local maximum) and the pole at the local minimum of $(Z_g + Z_v + Z_l)$. The simulations in Sec. II C show that each pole lies above its associated zero, and the distance in frequency between the zero and pole in each pair is correlated with the

glottal area, and thus inversely correlated with the magnitude of Z_g . The emergence of a zero-pole pair can be understood more clearly by considering the limiting case where $Z_l = \infty$, in which case Eq. (6) has a zero at each subglottal resonance. In sum, the net effect of adding subglottal coupling to a source-filter model is an extra zero-pole pair at each subglottal resonance.

B. Model calculations

Using the model developed in previous sections, a simulation was performed to show how the added zero (at the frequency of the second subglottal resonance) resulting from subglottal coupling and its associated pole lead to the frequency jump and amplitude attenuation effects seen in speech spectra. Taking account of radiation from the mouth, the pressure spectrum of an /a/ diphthong for a male speaker was simulated by changing dimensions linearly between the tube shapes approximating the vocal tract configurations of /a/ and /i/ shown in Fig. 4. The dimensions used were $A_b = 1 \text{ cm}^2$, $A_f = 8 \text{ cm}^2$, $l_b = 10.7 \text{ cm}$, $l_f = 9.0 \text{ cm}$ for /a/, and $A_b = 8 \text{ cm}^2$, $A_f = 1 \text{ cm}^2$, $l_b = 7.9 \text{ cm}$, $l_f = 6.9 \text{ cm}$ for /i/. The lossy capacitance Z_c was taken to be a resistance in series with a capacitance and another resistance in parallel. The values of these parameters were empirically chosen to approximate the subglottal spectrum near *SubF2*.

Using values within the ranges for a male speaker, we set $U_g = 200 \text{ cm}^3 \text{ s}^{-1}$ (Holmberg *et al.*, 1988), $l_g = 1.5 \text{ cm}$, $d_g = 0.045 \text{ cm}$, $h_g = 0.3 \text{ cm}$ (Li *et al.*, 2006), $M_g = 0.0082 \text{ g cm}^{-4}$, $R_g = 128.3 \text{ dyn s cm}^{-5}$ (Rösler and Strube, 1989; Stevens, 1998, p. 165), $l_l = 19.35 \text{ cm}$, $A_l = 2.5 \text{ cm}^2$ (Ishizaka *et al.*, 1976), and $A_g = 0.0675 \text{ cm}^2$ (Stevens, 1971). A_g , defined here as $l_g d_g$, is taken to represent the average value of the glottal area over one glottal cycle. With the understanding that the true glottal area is continuously changing, A_g will be referred to as “glottal area” for simplicity.

To simulate *F2* tracks, one last step was taken: a resistance of 1.5 acoustic ohms was added in series to Z_r to give the vocal tract formant peak some bandwidth. This was done as an alternative to adding wall impedances to facilitate finding modeled values of *F2*, since without wall impedances the vocal tract formant will always dominate any poles or zeroes arising from subglottal coupling. With this added resistance and with the above given parameter settings, the simulated track of the second formant prominence frequency matches one recorded by hand, using methods described in Sec. V A, for a male speaker’s /a/ utterance. This track is shown in Fig. 5, and does show a frequency jump effect near *SubF2*. Also shown in Fig. 5 is a simulated *F2* track using the same parameter values, but with no oral-subglottal coupling ($Z_g = \infty$). No jump occurs in this track, showing that the frequency jump observed in the coupled case arises directly from subglottal coupling, rather than this particular set of parameter values.

The frequency jump effect arises as the second formant peak passes through the zero pole pair introduced by subglottal coupling. The dynamics of this effect can be understood through the schematic pole-zero plot shown in Fig. 6. As

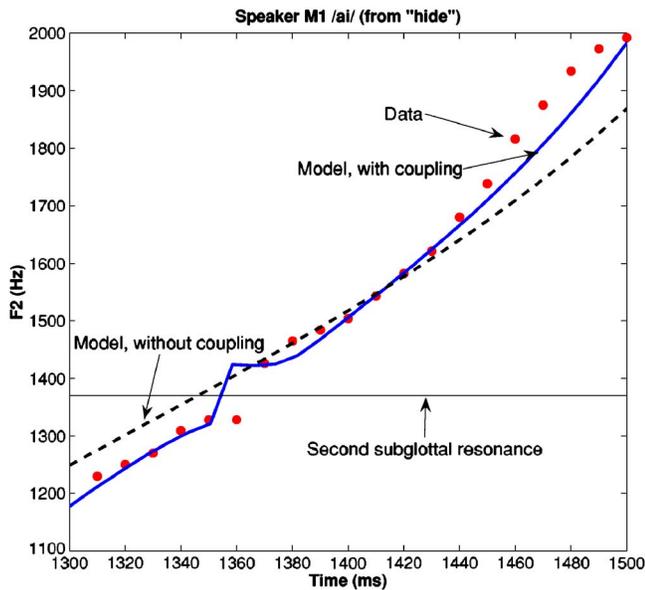


FIG. 5. (Color online) Observed and modeled F_2 tracks (with and without coupling) for an /ai/ diphthong spoken by a male speaker (M1). The modeled track with coupling shows a jump in F_2 near $SubF_2$, while the modeled track without coupling shows no discontinuities. Data were taken once per pitch period, with window length of one glottal cycle.

discussed in Sec. II A, the subglottal pole lies above the subglottal zero in frequency. As the pole associated with the second formant prominence approaches the zero located at $SubF_2$, the second formant peak amplitude is attenuated. The second formant peak then becomes associated with the higher pole and resumes its normal linear trajectory. The discontinuity seen in the F_2 track is a result of this switch in the formant peak's pole affiliation. A frequency discontinuity near $SubF_2$ thus arises directly from subglottal coupling in this model. As can be seen in Fig. 6, the size of this discontinuity will be approximately equal to the frequency difference between the subglottal zero and its associated pole.

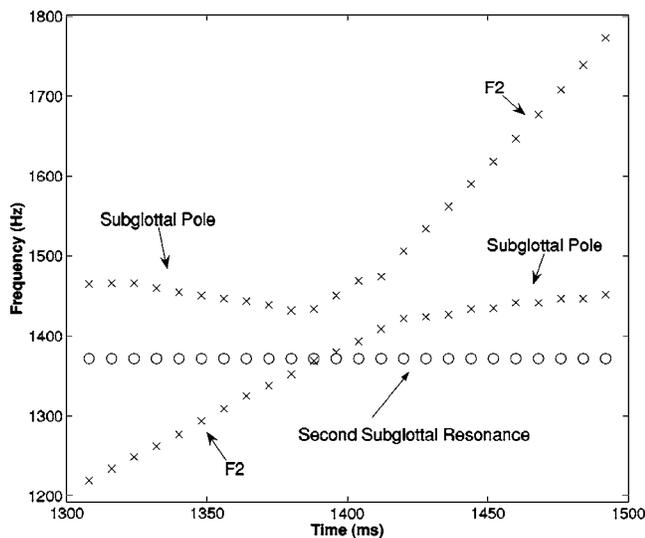


FIG. 6. Schematic of pole-zero plot corresponding to a simulated back-front diphthong with subglottal coupling. As the lower pole approaches the zero, the formant peak amplitude is attenuated and the peak frequency switches affiliation to the higher pole, which moves away from the zero.

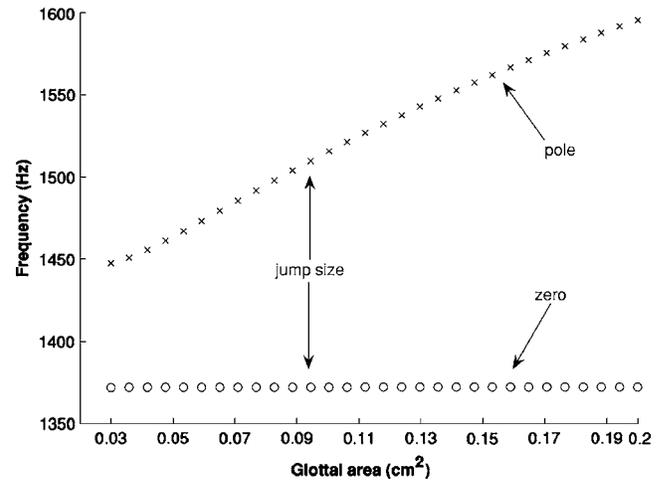


FIG. 7. Simulated frequencies of the zero-pole pair due to subglottal coupling near the second subglottal resonance frequency, where the zero is located. As glottal area increases from 0.03 to 0.2 cm^2 the zero-pole separation increases from 75 to 225 Hz, implying larger observed “jumps” in the second formant prominence frequency.

continuity will be approximately equal to the frequency difference between the subglottal zero and its associated pole.

C. Model predictions

This model predicts the occurrence of subglottal coupling effects; it also predicts patterns of cross-speaker variation in these effects. As discussed in Sec. II A, subglottal coupling increases as the magnitude of Z_g decreases, giving rise to greater separation between the pole and zero introduced by subglottal coupling. This leads to increased attenuation of the second formant prominence, A_2 (due to the more prominent zero, as the pole moves away), and an increased jump size in F_2 (due to the increased zero-pole distance). Because Z_g varies inversely with glottal area ($l_g d_g$) [Eq. (1)], a larger glottal opening implies greater subglottal coupling. Speakers with breathier voices (larger glottal area) are thus expected to show greater frequency jump and amplitude attenuation.

To examine the effect glottal area (and thus Z_g) has on the size of the frequency separation of the zero-pole pair of U_v/U_0 introduced by coupling (and thus predicted F_2 jump size), the pole and zero were calculated for a range of A_g . The influence of the vocal tract due to coupling was ignored by setting $Z_v=0$, so that the only variable is the average glottal area A_g . As A_g is increased from 0.03 to 0.2 cm^2 for normal modal voicing (Stevens, 1998, p. 35), predicted jump size, shown in Fig. 7, increases from 75 to 225 Hz as Z_g decreases, causing an increase in subglottal coupling and a resulting increase in the zero-pole distance.

To examine the effect glottal area (and thus Z_g), has on predicted A_2 attenuation near $SubF_2$, simulations of the vocal tract output for an /ai/ diphthong, as described in Sec. II B, were performed for a range of A_g . Increasing A_g from 0.03 to 0.2 cm^2 predicts an increase in A_2 attenuation from 5 to 13 dB, as shown in Fig. 8, with larger A_g causing greater attenuation near $SubF_2$. Because the model neglects wall impedances and uses a simple model for the subglottal imped-

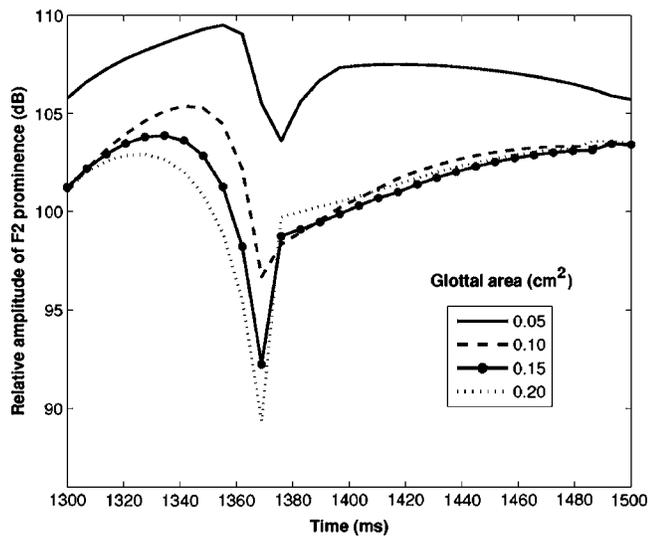


FIG. 8. Simulated F_2 amplitude trajectories for /aɪ/ diphthongs using glottal areas of 0.05, 0.10, 0.15, and 0.20 cm^2 . The amplitude attenuation seen in these trajectories occurs as the second formant peak crosses $SubF_2$ at around 1370 ms. When glottal area is 0.05 cm^2 , an attenuation of about 5 dB is observed in the F_2 prominence. For a larger glottal area of 0.20 cm^2 , the attenuation increases to 12 to 13 dB. Increased glottal area leads to greater zero-pole separation. The zero then becomes more prominent, causing greater amplitude attenuation to the F_2 prominence.

ance, these attenuation values are only approximate. These simplifications affect bandwidths, but not the general prediction of a positive correlation between glottal area and the strength of subglottal coupling effects.

Note that the 1.5 acoustic ohm added resistance described in Sec. II B was not used in the calculations for Fig. 7 because the vocal tract was decoupled, and was not used in Fig. 8 because it would affect measurement of the A2 dip near $SubF_2$.

Variation in the parameters used in this model will lead to variation in the extent of subglottal coupling effects for different speakers. The bulk of this paper shows the robustness of subglottal coupling effects across speakers and accounts for the differences observed in the strength of the coupling effects across speakers.

III. PROCEDURES

A. Data collection

Acoustic data were collected from three male and three female speakers, labeled M1–M3 and F1–F3. All speakers were native speakers of American English except M1, a native speaker of Canadian English. Subjects were recorded saying the phrase “hVd, say hVd again” five times for each of the American English monophthongs and diphthongs shown in Table I. The sentences were presented in a random order and prompted on a computer screen.

All utterances were recorded in a sound-isolated chamber using an Electro-Voice model DO54 omni-directional microphone (for audio data) and an accelerometer (for subglottal data). The microphone signal was first amplified by a mixer, then low-pass filtered at 4.8 kHz to avoid aliasing. It was digitized into a PC using the MARSHA³ program at a sampling rate of 10 kHz, which is above the Nyquist rate.

TABLE I. American English vowel IPA symbols, carrier words used, and [back] value.

Vowel	English word	[+/-back]
/i/	“heed”	–
/ɪ/	“hid”	–
/ɛ/	“head”	–
/æ/	“had”	–
/ɑ/	“hodd”	+
/ɔ/	“hawed”	+
/o/	“hoed”	+
/ʌ/	“hud”	+
/u/	“hood”	+
/ʊ/	“who’d”	+
/ɜ:/	“heard”	+
/aʊ/	“how’d”	+/+
/aɪ/	“hide”	+/-
/ɔɪ/	“hoid”	+/-
/ju/	“hued”	-/+
/eɪ/	“hade”	-/-

B. Measuring the subglottal resonances

The subglottal resonances for each speaker were measured using an EMkay BU-1771 accelerometer glued 1 in. above the sternal notch. This method is noninvasive and gives a spectrum sufficiently clear to accurately determine subglottal resonant frequencies (Henke, 1974; Stevens *et al.*, 1975; Cheyne, 2002). Because of its small mass (300 mg), the accelerometer does not have a significant impact on the system to which it is attached.

For a typical speaker, the accelerometer voltage signal ranges from 10 to 15 mVPP. A noninverting amplifier with a gain of 10 was built to amplify this voltage signal before it was fed to a mixer, and then digitized into a computer. A low offset, low drift LF411 operational amplifier was used to eliminate additional dc offset that came from the accelerometer output. The ac-coupled mixer removed the dc bias from the signal before it was digitized by the computer at 10 kHz. The digitized accelerometer signal was then bandpass filtered around the $SubF_2$ region to emphasize the second subglottal resonance, with a bandwidth of 600 Hz for male speakers and 800 Hz for female speakers.⁴ A wider bandwidth was used for female speakers due to their generally breathier voices, which give wider $SubF_2$ peaks (Klatt and Klatt, 1990; Hanson, 1997). A sample spectrogram and DFT of the accelerometer output for the word “heed” (spoken by speaker M1), taken at midvowel using a 10 ms window, are displayed in Fig. 9. The subglottal resonances for this utterance are at 550 and 1360 Hz.

IV. SECOND SUBGLOTTAL RESONANCE ANALYSIS

$SubF_2$ values from the accelerometer data were first found for all vowels for all speakers. Although only diphthongs will be examined in the remainder of this paper, $SubF_2$ values for both monophthongs and diphthongs were measured to explore the overall distribution of speakers’ $SubF_2$ values. Previous studies, listed in the following, have reported fewer than 10 $SubF_2$ measurements per speaker, on one to three different vowels.

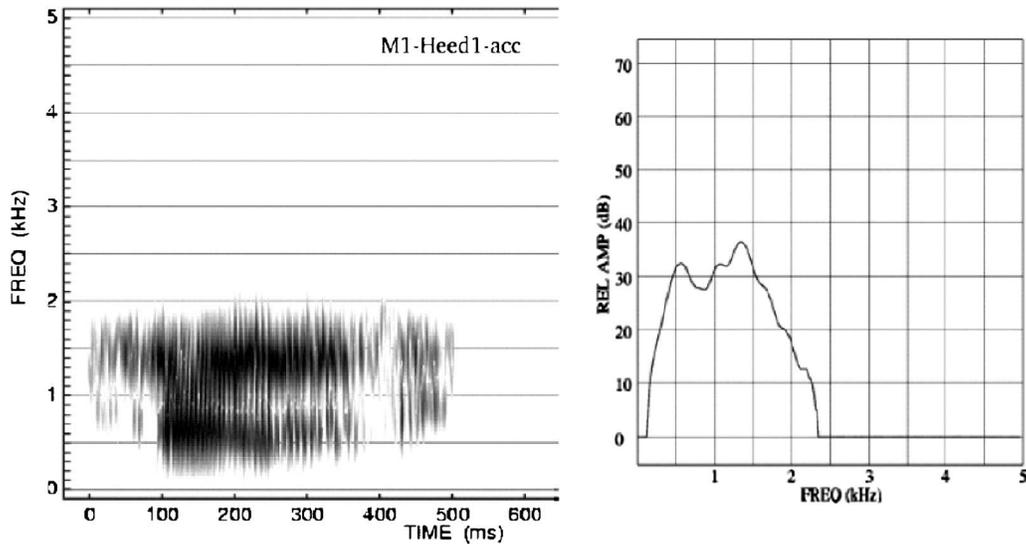


FIG. 9. Sample spectrogram and DFT of M1's accelerometer signal for "heed." The DFT was taken around 240 ms with a window size of 10 ms. Spectral peaks are visible at 550 Hz (*SubF1*) and 1360 Hz (*SubF2*). A small peak around 1000 Hz can be seen in the accelerometer spectrum. This peak is a result of coupling to the supraglottal tract, and is not captured by our simple model of the subglottal system. For a more detailed model accounting for such peaks, see Lulich (2006).

Because *SubF2* stays relatively constant during phonation, its values for all monophthongs were determined using a formant tracker. Means of data taken using the tracker were found to agree with hand-taken data for several test vowels and also with those measured by hand in the diphthong analysis in Sec. V. Statistics on each speaker's *SubF2* distribution are shown in Table II. Histograms of *SubF2* values for all speakers conform reasonably well to normal distributions and agree with previously reported *SubF2* values found using invasive techniques. Ishizaka *et al.* (1976) reported means "around 1400 Hz" for five speakers (male and female), and Cranen and Boves (1987) reported a mean of 1355 Hz for two male speakers, while mean values for this study's speakers fall between 1280 and 1620 Hz.⁵

V. DIPHTHONG ANALYSES

Back-front diphthongs of the six speakers were examined for coupling effects (formant frequency jump and amplitude attenuation) predicted by the theoretical model when the second formant peak passes *SubF2*. *F2*, second formant frequency, *A2*, amplitude of second formant prominence, and *SubF2*, second subglottal resonance, were measured by hand in two different ways in three steps. A Hamming window with length of one pitch period was first used to take mea-

surements from consecutive pitch periods for four diphthong tokens of each of the six speakers. Next, the window size was changed to slightly longer than half of each glottal period, and measurements on the same tokens were taken twice per pitch period to compare the open and closed phases of the glottal cycle. As open-phase measurements showed stronger evidence of subglottal coupling effects, open-phase measurements were finally made on six additional diphthong tokens for each speaker, totaling ten open-phase tokens per speaker.

A. Analysis by pitch period

Measurements described in Sec. V were taken by hand at consecutive pitch periods with a window size of one pitch period over the audio and accelerometer data for the diphthongs /aɪ/ ("hide") and /ɔɪ/ ("hoi"). Two repetitions of each diphthong were analyzed, for a total of four vowel tokens for each of the six speakers.

F2, *A2*, and *SubF2* time trajectories were plotted to examine *F2* and *A2* behavior near *SubF2*. An example is shown in Fig. 10. Beginning around 1320 ms, *F2* for speaker M2 jumps approximately 200 Hz as it passes through *SubF2*, while *A2* falls 6.7 dB, then rises 4 dB. The jump size falls within the range of 75–225 Hz predicted by the theoretical model and the attenuation falls within the predicted 5–13 dB range. Table III summarizes the average size of frequency jumps and *A2* attenuation for the six speakers. The size of the *F2* jump (0–249 Hz) and, to a lesser extent, the amount of *A2* attenuation (–8.5 to –6.4 dB; 5.6 to 8 dB) are speaker dependent because the distance between the zero-pole pair from subglottal coupling depends on an individual's glottal impedance, as described in Sec. II C.

As seen in Table III, an obvious jump in *F2* could not be observed for some tokens and speakers, but attenuation in *A2* near the frequency crossing was observed for all tokens and

TABLE II. Mean, standard deviation (σ), number of tokens (N), and normal distribution chi-squared for speakers' *SubF2* distributions, across monophthongs and diphthongs.

Speaker	\bar{x} (Hz)	σ (Hz)	N	χ^2_v
M1	1374	12	168	1.3
M2	1310	20	160	0.86
M3	1280	18	158	1.0
F1	1620	28	160	0.70
F2	1469	23	170	0.80
F3	1447	30	170	1.1

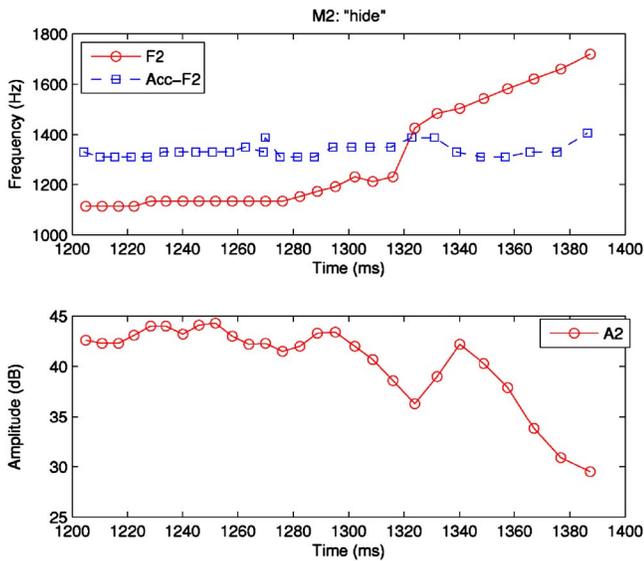


FIG. 10. (Color online) Time trajectories of $F2$, $SubF2$, and $A2$, measured with pitch-period window length, for a “hide” token spoken by a male speaker (M2). The $F2$ track shows a jump of 200 Hz and the $A2$ track shows a dip of -6.7 dB and rise of 4.0 dB at approximately 1320 ms.

speakers. In the above-mentioned diphthong analysis, tokens were marked “no jump” if a jump did occur, but did not cross $SubF2$. In tokens where this happened, the jump always occurred just above $SubF2$. This type of jump may be explainable in terms of the theoretical model. In the simulations of $F2$ tracks described in Sec. II B, if no resistance is added in series with Z_r , a frequency jump still occurs but generally takes place just above $SubF2$. Referring to Fig. 6, this corresponds to the lower pole having sufficiently low bandwidth and high amplitude to dominate the higher pole, even as the lower pole approaches $SubF2$. As the higher pole moves away from $SubF2$, the formant peak switches affiliation to the higher pole and a jump is observed above $SubF2$.

B. Open/closed phase analysis

A second method of analysis was used to determine whether $F2$ and $A2$ measurements would differ during parts

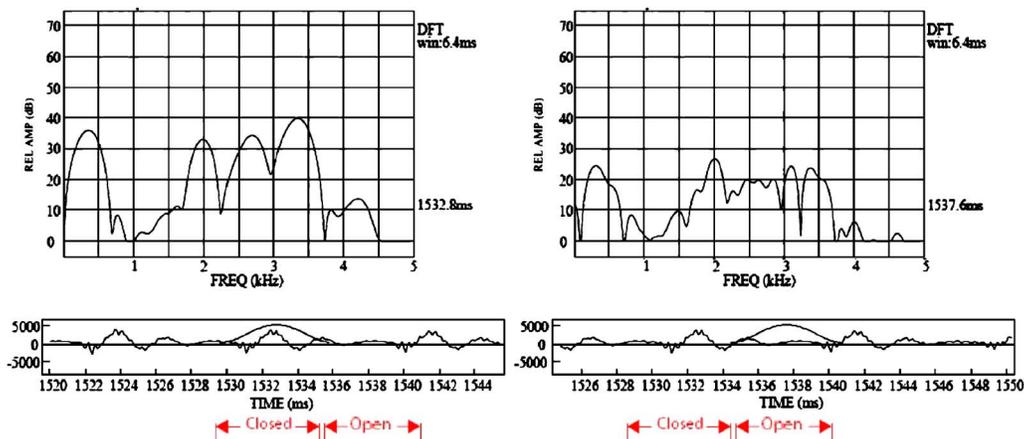


FIG. 11. (Color online) Spectra from two window placements within one pitch period of the vowel /i/. The left spectrum window is centered on the closed phase of the waveform. $F1$ and $F2$ are located at 350 and 2000 Hz. The right spectrum window is centered on the open phase of the waveform. Average spectral amplitude is lower during the open phase than during the closed phase.

TABLE III. $F2$ jump and $A2$ attenuation measurements (falling and rising values) for backfront diphthongs near $SubF2$, averaged across four tokens per speaker. Window size set at pitch period length. Data ranges are in parentheses.

Speaker	Average jump in $F2$ (Hz)	Average change in $A2$ at jump (dB)	
		Falling	Rising
M1	141 (78–195)	6.4(3.5–8.0)	7.4(3.5–10.0)
M2	220 (156–371)	8.5(5.4–13.6)	5.6(3.3–8.2)
M3	None	8.2(6.0–9.2)	8.0(6.0–8.9)
F1	None	7.0(5.0–8.0)	7.0(5.0–8.0)
F2	64 (0–254)	7.5(2.6–13.8)	6.5(2.1–12.8)
F3	249 (176–312)	6.7(3.4–10)	7.6(5.1–10)

of each pitch period when the glottis was more open or more closed. Since greater glottal area leads to greater coupling, one would expect greater coupling during the more open glottis phase than during the phase where the glottis is narrower. The window size was therefore adjusted to be slightly longer than one-half pitch period to better capture changes within each glottal cycle.

Measurements were taken by hand at two points within each pitch period on the same set of tokens as presented earlier, with window placement determined so that most of the windowed waveform would closely approximate either the more closed or more open portion of the glottal cycle under normal speech circumstances. Closed-phase measurements were taken with the window centered over the biggest peak in the pitch period, and open-phase measurements were taken with the window centered over the lowest amplitude part of the pitch period, before the largest peak in the next pitch period, as demonstrated in Fig. 11. It is recognized that the glottis may be closed for less than 50% of the glottal cycle, and that speakers with breathier voices might even never have complete glottal closure, but “open phase” and “closed phase” remain useful idealizations, to be understood here as the portion of a pitch period where the glottis is more open and more closed.

In attempting to track the effect of $SubF2$ on the $F2$ and $A2$ tracks, a very small Hamming window was used to select

either the open or closed phase of the glottal cycle. Such a short time window corresponds to poor resolution in the frequency domain. According to Rayleigh’s criterion, the minimum distance between two frequency peaks that can be resolved by a window is half of the window’s main lobe width. Hamming windows have a main-lobe width of $8\pi/N$, where N is the number of points of the window. Given window time duration t , the main-lobe width is found to be $w=4/t$ (Hz) (Oppenheim and Schaffer, 1999, p. 471). A 4 ms Hamming window at a rate of 10 kHz thus has a width of 1000 Hz, and theoretically is only able to resolve peaks that are 500 Hz apart. However, it was empirically found that peaks separated by 400 Hz could also be resolved. By the same calculation, a 6.4 ms Hamming window would resolve two peaks that are 312 Hz apart. This frequency resolution is still too poor to resolve two peaks that correspond to $F2$ and the pole introduced by trachea coupling, when they come too close together. Given such a small window, it can only be concluded that the single $F2$ peak in the audio spectrum is the smoothed result of two peaks that are close together. Because the window itself is symmetric, the peak of the windowed signal would lie closer in frequency to the peak that has a higher amplitude. One can thus conclude that the frequency jump size given by the measured $F2$ is actually smaller than or equal to the actual frequency jump, making the observed frequency jump a conservative estimate.

To track the effect of $SubF2$ on $F2$ and $A2$ during the open and closed phases of the glottal cycle, $F2$ values recorded (from the audio data) during the closed phase ($F2_{closed}$), and during the open phase ($F2_{open}$), together with the second subglottal frequency (from the accelerometer data) taken during the closed phase ($SubF2_{closed}$), were all plotted against time. $SubF2_{closed}$ was used because $SubF2$ should be less affected by subglottal coupling when the glottis is more closed, and thus more accurate. The example in Fig. 12 illustrates how measurements taken during the two phases can differ. In this example, a frequency jump near $SubF2$ and a corresponding amplitude dip occur in open-phase measurements, but not in closed-phase measurements. $A2_{open}$ is also consistently lower than $A2_{closed}$ in Fig. 12, as expected because there is more energy during the initial glottal closure. From examination of the decaying waveform within a pitch period, it can be seen that the open-phase signal amplitude is approximately 9 dB lower than the closed-phase signal amplitude.⁶

Table IV summarizes the size of observed frequency jumps and $A2$ attenuation during the open and closed phases. For all open-phase and most closed-phase measurements, $A2$ attenuation occurs when $F2$ comes close to $SubF2$, in agreement with the results in Sec. V A. Little difference was found in the amount of attenuation between open and closed-phase measurements, but the only instances where attenuation did not occur were in a few closed-phase measurements, and in this sense attenuation was slightly more prevalent in open-phase measurements. The size of frequency jumps is generally greater for open-phase measurements, as expected if frequency jumps arise from subglottal coupling. Frequency jumps can still occur during closed-phase measurements because the size of the glottal opening within each period is

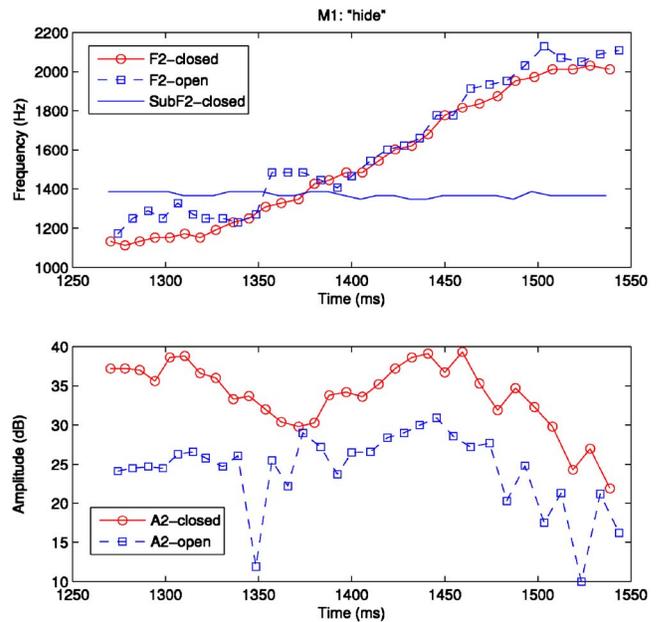


FIG. 12. (Color online) Time trajectories of $F2_{closed}$, $F2_{open}$, $SubF2_{closed}$, $A2_{closed}$, $A2_{open}$ for a version of “hide” produced by a female speaker (F1). In the top panel, the curve with circles represents closed-phase $F2$. Assuming the glottis is completely closed, no interference on $F2$ from subglottal coupling near $SubF2$ should be seen. The closed-phase curve rises smoothly during the /ai/ diphthong, as expected. The open-phase $F2$ trajectory, by contrast, is greatly affected by the zero-pole pair from subglottal coupling, resulting in a frequency jump (of 210 Hz) near $SubF2$. The lower panel shows the corresponding amplitude trajectories. The closed-phase $A2$ trajectory has greater amplitude due to the presence of more energy in the initial closing of the glottis, while the open-phase trajectory shows an amplitude dip (of 14 dB) near the frequency jump, due to attenuation near the zero introduced by subglottal coupling.

highly speaker-dependent, and some speakers might maintain a glottal chink throughout the entire utterance.

The comparison of open and closed-phase measurements confirms that open-phase measurements are better indicators of the effect of subglottal coupling on formant frequencies, and to some extent on amplitudes. Open-phase measurements were thus taken for six additional diphthong tokens per speaker (totaling five “hide” and five “hoid” tokens per speaker). Table V lists average open-phase $F2$ jump and $A2$ attenuation values for each of the six speakers.

C. Discussion of diphthong results

The results of the diphthong analysis in Tables III–V show that as the second formant crosses $SubF2$, there is often

TABLE IV. Amount of $F2$ jump and $A2$ attenuation in back-front diphthongs due to sub-glottal coupling, measured during open and closed phases of the glottal cycle. Averaged across four tokens for each speaker. Data ranges are in parentheses.

Speaker	Average $F2$ jump (Hz)		Average $A2$ attenuation (dB)	
	Closed phase	Open phase	Closed phase	Open phase
M1	141 (74–176)	224(176–273)	7.1 (3–14)	12.4 (5–22)
M2	229(137–350)	220 (97–274)	6.4 (2–12)	8.9 (4–18)
M3	103 (0–410)	196 (78–313)	6.8 (0–13)	5.2 (0.5–9)
F1	68 (0–273)	327(234–390)	6.3 (0–17)	7.4 (1–15)
F2	49 (0–196)	171(137–215)	10.4(1.5–17)	6.5 (1–12)
F3	103 (0–410)	131 (0–349)	8.6(2.1–16)	8.0(1.5–15)

TABLE V. Open-phase $F2$ jumps and $A2$ attenuation (falling and rising values) in diphthongs near $SubF2$, averaged over ten tokens (5 /ai/ and 5 /oi/). The correlation between each speaker's jump size and $A2$ attenuation amount (average of falling and rising values) and the correlation significances are shown in the rightmost columns.

Speaker	Average Jump in $F2$ (Hz)	Average $\Delta A2$ at jump (dB)	Corr. (r)	Corr. significance
M1	299	-8.3, +7.5	-0.46	$p=0.18$
M2	348	-8.0, +6.5	-0.45	$p=0.19$
M3	237	-5.6, +6.5	0.58	$p=0.08$
F1	294	-8.5, +9.4	-0.53	$p=0.12$
F2	252	-8.4, +6.2	0.05	$p=0.89$
F3	201	-6.2, +8.3	-0.60	$p=0.07$

a discontinuity in the $F2$ track due to subglottal coupling. A more robust cue for this coupling is $A2$ attenuation near $SubF2$, which occurs regardless of whether there is an observable frequency jump. The ranges for the frequency jump and amplitude attenuation vary among speakers due to differences in glottal impedance.

Frequency jumps were found to be around 50–300 Hz, and most data fell within the predicted jump size range of 75–225 Hz. However, some jump size measurements fell above the predicted range, a discrepancy which could be due to the lack of finite, frequency-dependent vocal tract wall impedances in simulating jump sizes in Fig. 7. Finite, frequency-dependent wall impedances would increase the simulated formant's bandwidth and decrease its amplitude, causing it to be canceled by the subglottal zero sooner and to reemerge once its pole affiliation is switched later, leading to larger observed frequency discontinuities between these two events.

Some diphthong tokens showed no frequency jump. This may occur because of variation in the bandwidths, amplitudes, and separation of the subglottal pole and zero by speaker and utterance: Unless the pole and zero are sufficiently sharp and separated, their effect on the spectrum will be indistinguishable from noise. By this argument, only frequency jumps above some lower limit should be observed; otherwise no jump will be observed. This is consistent with the diphthong data, where very few jumps of less than 100 Hz were observed.

While the 5–12 dB $A2$ attenuation observed in the analyses agrees well with the predicted range of 5–13 dB, using a more accurate subglottal system model that incorporates bronchi structure (see Ishizaka *et al.*, 1976; Lulich, 2006) and adding wall impedances to the vocal tract would give understanding of how model parameter values affect the size of subglottal coupling effects.

A weak negative correlation (though not always statistically significant) exists between a speaker's $F2$ jump size and the corresponding $A2$ attenuation for four of the six speakers. This can be explained by the velocity with which the formant moves across the region of the spectrum containing the zero-pole pair introduced by subglottal coupling. Given the fixed number of measurement points at every glottal cycle, the faster the formant passes through the zero-pole region in time, the larger the observed frequency jump size,

and the smaller the observed $A2$ attenuation due to the short amount of time the formant is affected by subglottal coupling in the zero-pole region.

VI. IMPLICATIONS FOR QUANTAL THEORY

This paper's findings can be seen as part of the larger project of testing QT directly by testing a proposed quantal relation. One aspect of QT predicts that speakers will avoid putting formants in an acoustically unstable region, which will then be used as a dividing line between sounds having + or - values of a phonological feature. Showing that this acoustic-articulatory nonlinearity occurs robustly across speakers is a necessary condition, and thus the first step, for testing whether it is used as a quantal relation. The diphthong analyses carried out earlier show that subglottal coupling effects occurred robustly for all six speakers. In another study, Lulich (2006) observed the same unstable region in a data set of 400 CV transitions from one English speaker. Both findings suggest that the frequency range near $SubF2$ forms an acoustically unstable region which could be used by speakers to separate front and back vowels.

The next step, testing whether the acoustically unstable region is used in this way, requires study of vowel production and perception near $SubF2$ and comparison of the quantal hypothesis to alternative hypotheses. This step is not addressed in this paper's analysis, but has been in recent work. Sonderegger (2004) used a larger version of the current data set, containing monophthong (shown in Table I) and $SubF2$ data for 14 English speakers, to examine whether $F2$ fell below or above $SubF2$ for all monophthongs for all speakers. $F2$ was always significantly ($p < 0.05$) above $SubF2$ for front vowels and almost always significantly below ($p < 0.05$) $SubF2$ for back vowels, showing a pattern which supports the quantal hypothesis. On the perception side, a recent study (Lulich *et al.*, 2007; Lulich, 2006) analyzed the effect of $SubF2$ location on the categorization of vowel tokens as [+back] or [-back] by varying the frequency of the subglottal zero in synthesized vowel tokens. The data supported the hypothesis that speakers use the location of $SubF2$ (manifested as a zero in the vowel spectrum) to distinguish front and back vowels. In addition, the quantal hypothesis was tested against several alternative hypotheses, and was found to be more consistent with the data.

An important but more involved part of testing a quantal relation is studying whether it holds cross-linguistically. This study and those mentioned earlier all focus only on English vowels. However, quantal relations are hypothesized to be language independent, because speakers of different languages should have the same nonlinear acoustic-articulatory relations.

While more work is needed before subglottal coupling effects near $SubF2$ can be said to be "tested" as a quantal relation for [back], the current paper, together with the above-described studies, support the hypothesis that (1) subglottal coupling effects near $SubF2$ lead to an acoustically unstable region and (2) these coupling effects form a quantal relation dividing front and back vowels, at least in English and predicted cross-linguistically.

VII. CONCLUSION

Accelerometer measurements of the second subglottal resonance agreed well with previously reported values from invasive studies, confirming that an accelerometer glued below the glottal area can be used for noninvasive and relatively accurate measurements of subglottal resonances. When averaged across all tokens for all American English monophthongs and diphthongs, male speakers had mean *SubF2* values ranging from 1280 to 1374 Hz, female speakers from 1447 to 1620 Hz.

The effect of oral-subglottal coupling on the second formant was predicted using a theoretical model and simulation of *A2* attenuation and *F2* jumps near *SubF2* in back-front diphthongs. For all speakers, the back-front diphthongs /a/ and /ɔ/ consistently showed *A2* attenuation of 5–12 dB when *F2* approached *SubF2*, often accompanied by a frequency jump in the range of 50–300 Hz in *F2*; otherwise no jump occurred. Attenuation was present in nearly all diphthongs, even when jumps were not. Jumps (when present) were within or above the range predicted by the theoretical model and the predicted and observed ranges of attenuation were almost exactly the same, so that the subglottal coupling effects observed were in a sense stronger than predicted. Both attenuation and jumps were generally stronger in measurements taken during the open phase of the glottal cycle than during the closed phase, as expected if these effects result from subglottal coupling.

The coupling effects examined have potential relevance for speech synthesis and recognition systems, as well as further research on the role of subglottal resonances in speech. Inserting disturbances due to oral-subglottal coupling into the open phase of a pitch period could enhance the naturalness of synthesized speech. Subglottal coupling effects near the first and third subglottal resonances could also be examined, both for general understanding and for possible quantal effects involving the first and third vowel formants. A recent dissertation (Lulich, 2006) uses acoustic models and perception experiments to examine the role of lower airway resonances in defining distinctive feature contrasts via the zero-pole pairs they contribute to the transfer function.

The procedure used in this paper to examine the second subglottal resonance and its effects on the front-back division in vowel space forms part of a general methodology for studying articulatory-acoustic nonlinearities and testing whether they serve as quantal relations for phonological features. First, an acoustic model was used to explain how the hypothesized nonlinear effects (in this case attenuation effects and frequency jump near *SubF2*) could occur. Second, data were examined in detail to check whether the predicted effect occurred robustly across speakers. This lays the basis for studying whether the robust effect is indeed quantal, in this case separating front from back vowels. More generally, this method offers a way of investigating articulatory-acoustic nonlinearities which constrain the possible acoustics of linguistic sounds. This paper aims to show the value of investigating one such nonlinearity from acoustic phonetic and linguistic perspectives; its methodology could be used in further research to understand the many nonlinearities found

in human speech and their practical implications for speech research.

ACKNOWLEDGMENTS

First and foremost, we would like to thank Professor Ken Stevens for his knowledge, support, and advice. We also thank John Rodriguez, Harold Cheyne, Martin Hughes, John Sweeney, Arlin Mason, and members of the Speech Communication group for help with the experiment, as well as our subjects. We are grateful to Steven Lulich, Elisabeth Hon, and Yoko Saikachi for comments and discussion. This work was supported in part by NIH Grant No. DC00075 and the MIT UROP Program.

¹The measurement designated *F2* refers to the frequency of the spectrum peak corresponding to the second formant, which may deviate slightly from the actual natural frequency of the formant because of the influence of the zero arising from coupling to the subglottal tract. Likewise *A2* is the amplitude of the spectrum peak that is labeled as *F2*.

²The tube's length and impedances can be adjusted to fit an empirically measured subglottal spectrum peak's center frequency and bandwidth.

³Written by Mark Tiede of the MIT Speech Communication Group. MARSHA allows the user to record and digitize multiple channels of input at once (in this case the audio and accelerometer signals).

⁴The bandpass filter was used to emphasize the frequency range around *SubF2* mainly because the frequency range below *SubF2* has much higher energy. This is why there is still some low-frequency energy present after filtering.

⁵Independent of the measurements of *SubF2* taken from the subglottal tracks, it was found that one quick way to approximate a speaker's second subglottal resonance without the aid of an accelerometer is by inspecting the spectrum of a front vowel such as /i/, where *F2* is located far above *SubF2*. The pole from the zero-pole pair due to oral-subglottal coupling is visible in the speech spectrum and lies above *SubF2*. Its distance from *SubF2* is dependent on the glottal impedance, Z_g .

⁶For speaker F1, the bandwidth of the first formant, *B1*, is approximately 160 Hz; $B1 = 1/\tau\pi$, where τ is the time constant of the decaying waveform envelope. In this case $\tau = 0.002$ s, the time it takes for the amplitude of the wave to decrease by a factor of e , or 8.69 dB.

Cheyne, H. (2002). "Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Cranen, B., and Boves, L. (1987). "On subglottal formant analysis," *J. Acoust. Soc. Am.* **84**, 734–746.

Fant, G., Ishizaka, K., Lindqvist, J., and Sundberg, J. (1972). "Subglottal formants," *KTH Speech Transmission Laboratory Quarterly Progress and Status Report* 1, 1–12.

Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Hanson, H. (1997). "Glottal characteristics of female speakers: Acoustic correlates," *J. Acoust. Soc. Am.* **101**, 466–481.

Hanson, H., and Stevens, K. N. (1995). "Sub-glottal resonances in female speakers and their effect on vowel spectra," *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, vol. 3, pp. 182–185.

Henke, W. L. (1974). "Signals from external accelerometers during phonation: Attributes and their internal physical correlates," Technical Report, Massachusetts Institute of Technology, RLE Cambridge, MA.

Holmberg, E., Hillman, R., and Perkell, J. (1988). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *J. Acoust. Soc. Am.* **84**, 511–529.

Ishizaka, K., Matsudaira, M., and Kaneki, T. (1976). "Input acoustic-impedance measurement of the subglottal system," *J. Acoust. Soc. Am.* **60**, 190–197.

Kinsler, L., and Frey, A. (1962). *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York).

Klatt, D., and Klatt, L. (1990). "Analysis, synthesis and perception of voice quality variations among male and female talkers," *J. Acoust. Soc. Am.* **87**, 820–856.

- Ladefoged, P., and Bhaskararao, P. (1983). "Non-quantal aspects of consonant production: A study of retroflex consonants," *J. Phonetics* **11**, 291–302.
- Li, S., Scherer, R., Wan, M., and Wang, S. (2006). "The effect of three-dimensional glottal geometry on intraglottal quasi-steady flow distributions and their relationship with phonation," *Sci. China, Ser. C: Life Sci.* **49**, 82–88.
- Lulich, S. (2006). "The role of lower airway resonances in defining feature contrasts," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Lulich, S., Bachrach, A., and Malyska, N. (2007). *J. Acoust. Soc. Am.* (in press).
- Oppenheim, A., and Schaffer, R. (1999). *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).
- Perkell, J., and Cohen, M. (1989). "An indirect test of the quantal nature of speech in the production of the vowels /i/, /a/ and /u/," *J. Phonetics* **17**, 123–133.
- Pisoni, D. (1981). "Variability and the quantal theory of speech: A first report," *Phonetica* **37**, 285–305.
- Rösler, S., and Strube, H. (1989). "Measurement of the glottal impedance with a mechanical model," *J. Acoust. Soc. Am.* **86**, 1708–1716.
- Sonderegger, M. (2004). "Subglottal coupling and vowel space: An investigation in quantal theory," Physics B.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Stevens, K. (1971). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," *J. Acoust. Soc. Am.* **50**, 1180–1192.
- Stevens, K. (1972). "The quantal nature of speech: Evidence from articulatory-acoustic data," in *Human Communication: A Unified View*, edited by E. David and P. Denes, (McGraw Hill, New York), pp. 51–66.
- Stevens, K. (1989). "On the quantal nature of speech," *J. Phonetics* **17**, 3–46.
- Stevens, K. (1998). *Acoustic Phonetics* (MIT, Cambridge, MA).
- Stevens, K., and Blumstein, S. (1975). "Quantal aspects of consonant production and perception: A study of retroflex stop consonants," *J. Phonetics* **3**, 215–233.
- Stevens, K., Kalikow, D., and Willemain, T. (1975). "A miniature accelerometer for detecting glottal waveforms and nasalisation," *J. Speech Hear. Res.* **18**, 594–599.
- Zhao, W., Zhang, C., Frankel, S., and Mongeau, L. (2002). "Computational aeroacoustics of phonation. I. Computational methods and sound generation mechanisms," *J. Acoust. Soc. Am.* **112**, 2147–2154.