Multidimensional analyses of voicing offsets and onsets in female speakers[a)]

Laura L. Koenig[b)]
Haskins Laboratories, New Haven, Connecticut 06511, and Long Island University, Brooklyn, New York

W. Einar Mencl[c)]
Haskins Laboratories, New Haven, Connecticut 06511

Jorge C. Lucero[d)]
Department of Mathematics, University of Brasilia, Brazil

(Received 24 August 2004; revised 20 July 2005; accepted 20 July 2005)

This study investigates cross-speaker differences in the factors that predict voicing thresholds during abduction–adduction gestures in six normal women. Measures of baseline airflow, pulse amplitude, subglottal pressure, and fundamental frequency were made at voicing offset and onset during intervocalic /h/, produced in varying vowel environments and at different loudness levels, and subjected to relational analyses to determine which factors were most strongly related to the timing of voicing cessation or initiation. The data indicate that (a) all speakers showed differences between voicing offsets and onsets, but the degree of this effect varied across speakers; (b) loudness and vowel environment have speaker-specific effects on the likelihood of devoicing during /h/; and (c) baseline flow measures significantly predicted times of voicing offset and onset in all participants, but other variables contributing to voice timing differed across speakers. Overall, the results suggest that individual speakers have unique methods of achieving phonatory goals during running speech. These data contribute to the literature on individual differences in laryngeal function, and serve as a means of evaluating how well laryngeal models can reproduce the range of voicing behavior used by speakers during running speech tasks. © *2005 Acoustical Society of America.*
[DOI: 10.1121/1.2033572]

PACS number(s): 43.70.Aj, 43.70.Gr [ARB] Pages: 2535–2550

I. INTRODUCTION

The goal of this work is to investigate cross-speaker variation in the conditions under which sustained vocal-fold vibration (voicing) ceases and begins again during abduction–adduction gestures in running speech. Our measurements, made on noninvasively obtained aerodynamic signals from six women, sample across the range of factors that are known to affect voicing thresholds. Our analyses are designed to determine the factors most responsible for sustaining and/or achieving voicing in individual speakers.

Since van den Berg (1958), researchers have recognized that voicing depends upon achieving a balance among vocal-fold parameters and aerodynamic forces. Vocal-fold parameters include thickness, degree of adduction, glottal shape, and tissue characteristics (longitudinal tension, damping); aerodynamic forces include the transglottal pressure and the Bernoulli effect. In recent decades, theoretical and modeling studies have formalized the requirements for voicing in laryngeal models of varying complexity, capturing many features of human vocal behavior and providing estimates of the quantities involved in achieving and sustaining phonation (e.g., Ishizaka and Flanagan, 1972; Stevens, 1977; Titze, 1988, 1989, 1992).

Yet, relatively few data exist on the aerodynamic and laryngeal conditions at voicing offset and onset in living humans performing connected speech tasks involving abduction as well as adduction. Empirical investigations of phonation have frequently relied on measurements made from nonhuman (typically, canine) larynges (e.g., Alipour *et al.*, 1997; Alipour-Haghighi and Titze, 1991; Berke *et al.*, 1989; Finkelhor *et al.*, 1987; Saito *et al.*, 1983; Titze *et al.*, 1993; Yumoto *et al.*, 1993) or from excised human larynges (e.g., Baer, 1975; Matsushita, 1975; van den Berg and Tan, 1959). Such experiments elucidate general principles of vocal-fold vibration, but there are also limitations: Laryngeal geometry and histology differ across species (Cox *et al.*, 1999; Jiang *et al.*, 2001; Kim *et al.*, 2004), and excised larynges lack natural patterns of muscle contraction. The position of the human larynx is such that direct measurement of phonation during running speech requires using patients who have had surgical alterations such as tracheostomy or hemilaryngectomy (e.g., Hirano *et al.*, 1991, Jiang and Titze, 1993) or performing invasive techniques including tracheal puncture and/or inserting cameras and light sources into the pharynx (e.g., Baer *et al.*, 1983; Hertegård and Gauffin, 1995; Hertegård *et al.*, 1995; Plant *et al.*, 2004; Timcke *et al.*, 1958). Investigations using such methods typically record data from few speakers

[a)]Portions of this work were presented at the 3rd and 4th International Conferences on Voice Physiology and Biomechanics (Denver, September, 2002, and Marseille, August, 2004).
[b)]Author to whom correspondence should be addressed. Electronic mail: koenig@haskins.yale.edu
[c)]Electronic mail: einar@haskins.yale.edu
[d)]Electronic mail: lucero@mat.unb.br

(often, only one), making it difficult to ascertain the degree of cross-speaker variation. Less invasive observation of laryngeal function via the oral cavity (e.g., Kitzing and Sonneson, 1974) permits recording a larger number of speakers, but imposes strict limitations on the speech task (usually to sustained low vowels). To complement this work, many researchers have relied on indirect measurements made on acoustic, electroglottographic, or oral airflow signals (e.g., Behrman and Baken, 1997; Childers and Lee, 1991; Dromey *et al.*, 1992, Hanson, 1997; Higgins and Saxman, 1993; Holmberg *et al.*, 1988, 1989, 1994; Löfqvist *et al.*, 1995; Ní Chasaide and Gobl, 1993; Price, 1989; Rothenberg, 1973). These methods facilitate measurement of multiple speakers, and also allow for recording protected populations such as children. These studies have provided much information on phonatory behavior in steady-state (adducted) conditions, again contributing to our general understanding of vocal-fold vibration. With a few exceptions (Löfqvist *et al.*, 1995; Löfqvist and McGowan, 1992; Ní Chasaide and Gobl, 1993), however, this literature has typically considered sustained vowels or the vocalic portions of repetitive CV syllable productions. In naturalistic, running speech, the larynx alternates between adducted and abducted postures. Thus, to understand fully how speakers control phonation during speech, we also need to consider voicing characteristics in the vicinity of an abduction gesture.

Past work on phonation threshold pressures (P_{th}; Ishizaka and Flanagan, 1972; Titze, 1992) provides insight into the factors that speakers can manipulate to achieve phonation offsets and onsets. Variations in P_{th} reflect changes in the conditions that make phonation more or less likely. For the body-cover model of the vocal folds (Titze, 1988), P_{th} is defined as follows:

$$P_{\mathrm{th}} = \frac{cBx_0k_t}{T}, \tag{1}$$

where c=the mucosal wave velocity in the vocal-fold cover; B=tissue damping; x_0=glottal half-width; k_t is a translaryngeal pressure coefficient; and T is the thickness of the vocal folds. From this equation, we infer that the likelihood of phonation will vary with changes in translaryngeal pressure, abduction degree, longitudinal tension of the vocal folds (which affects the tissue compliance and, consequently, the mucosal wave velocity; cf. Titze, 1992), tissue damping, and vocal-fold thickness. Both tissue damping and vocal-fold thickness may vary with laryngeal setting.

Since the parameters that affect phonation thresholds are relatively independent of each other, it follows that speakers can satisfy the physical requirements for phonation offset and onset in a variety of ways. One question raised in past work is whether speakers actively increase longitudinal tension of the vocal folds to suppress voicing during speech. Halle and Stevens (1971) proposed that voiceless consonants such as [p, pʰ, h] are characterized by stiff vocal folds, whereas voiced consonants such as [b, ɦ] have slack vocal folds. Authors investigating laryngeal muscle activity in the vicinity of voiced and voiceless consonants (e.g., Dixit and MacNeilage, 1980; Kagaya and Hirose, 1975; Löfqvist *et al.*, 1989) have obtained conflicting findings. One possible explanation for the disagreement across studies is that individuals vary in whether or not they supplement abduction with increased vocal-fold tension. Cross-speaker variation in voicing and devoicing strategies would be consistent with the larger literature on speech production, where researchers have observed that different speakers may produce the same consonant or vowel with different muscle activation patterns (Raphael and Bell-Berti, 1975) and/or articulatory postures (Borden and Gay, 1979; Johnson *et al.*, 1993). Although some studies of glottal function have described speaker differences (e.g., Hanson, 1997; Löfqvist *et al.*, 1995), these have not explicitly considered cross-speaker variability in the presence or absence of phonation, or in the methods speakers use to control voicing offsets and onsets.

The focus of our analysis is specifically on the factors that relate to the *timing* of phonation offset and onset. This emphasis is based on the extensive literature indicating that voice timing differentiates phonologically voiced and voiceless consonants across diverse languages (e.g., Abramson and Lisker, 1970, 1985; Lisker and Abramson, 1964, 1970). Whereas most of the work on contrastive voicing has investigated obstruent consonants, our work investigates phonation offsets and onsets in the context of /h/. Since /h/ involves abduction in the context of a relatively open vocal tract, it provides a convenient method of assessing phonatory behavior via easily obtained oral airflow signals. It is true that supraglottal conditions differ between /h/ and oral obstruents; in particular, supraglottal pressure does not vary much during the production of /h/, whereas obstruents involve a pressure buildup in the supraglottal vocal tract, which affects the transglottal pressure differential and, possibly, aspects of glottal geometry (Bickley and Stevens, 1986). Despite these differences, there is evidence that analysis of /h/ can provide insight into phonatory behavior for other voiceless consonants. First, most speakers systematically produce /h/ with a vocal-fold abduction gesture (e.g., Koenig, 2000; Klatt *et al.*, 1968; Löfqvist *et al.*, 1995), and measures of voice source features surrounding /h/ are often qualitatively similar to those features measured before and after other voiceless consonants (Löfqvist *et al.*, 1995). Further, within speakers, the voicing characteristics of /h/ and those of contrastively voiceless stops show significant correlations (Koenig, 2000). Thus, laryngeal behavior in /h/ can serve as a foundation for understanding voicing control in consonants for which voicing is contrastive. For comparison with the results of modeling, /h/ data are also useful in that we can consider phonatory behavior free of the complicating effects of upper vocal-tract constrictions.

As part of characterizing phonation offset and onsets, we consider how the two conditions differ. Several past studies have indicated that the requirements for initiating phonation are more stringent than those for sustaining it (e.g., Baer, 1975; Berry *et al.*, 1995, Chan *et al.*, 1997, Hirose and Niimi, 1987; Lindqvist, 1972; Lisker *et al.*, 1970; Munhall *et al.*, 1994, Titze *et al.*, 1995). The physical principles underlying this hysteresis phenomenon have been described in some detail for simple laryngeal models (e.g., Lucero, 1995, 1999). By quantifying the differences between voicing off-

TABLE I. Speaker information.

Speaker	Age (years)	Grew up where?	Devoiced /h/ (%)
F1	32	ME	52.3
F2	22	IA	37.1
F3	22	CT	28.1
F4	26	NY	56.3
F5	34	NY	54.8
F6	28	NY	85.9

sets and onsets in our data, we hope to inform future studies of hysteresis using more complex laryngeal models.

In general, then, we seek to provide detailed data on phonatory behavior in multiple speakers, to elucidate the range of strategies that individuals use to achieve phonatory goals, and to serve as a reference for evaluating how accurately laryngeal models reproduce the voicing patterns of normal speakers performing connected speech tasks. The data presented here are drawn from a larger study comparing voicing behavior for /h/ in several male and female speakers. Theoretical considerations and our past work (Koenig, 2000; Lucero and Koenig, 2005) suggest that women are more likely to devoice during an abduction gesture than men are. In this paper, we focus on female speakers who devoice at least 25% of their /h/ productions. In future manuscripts, we will undertake comparisons between men and women, and between speakers who produce mostly voiced /h/ vs those who produce a mixture of voiced and voiceless /h/.

II. METHODS

A. Speakers

Data were obtained from six adult females, who provided informed consent to take part in the experiment. All spoke American English as their native language; none had a strong regional accent. Participants were nonsmoking, in good health at the time of recording, reported no history of speech-, language-, or hearing disorders, and had vocal qualities within normal limits as judged informally by the first author. To restrict variation in vocal parameters as a function of aging, speakers were required to be between 20 and 40 years of age. Studies of laryngeal changes across the life span (Hirano *et al.*, 1983; Kahane, 1987, 1988) show little change in either the vocal folds or the laryngeal cartilages over this age range. Since the present analysis sought to quantify the conditions around voicing offset and onset, a final selection criterion was that speakers show a voicing break in 25% or more of their /h/ productions, to ensure that at least 50 tokens were available for statistical comparison of voicing offsets vs onsets. Table I shows speaker information and the percentage of devoiced /h/ produced by each speaker.

B. Speech materials

Speakers were recorded producing the following utterances, in three blocks of normal, loud, and soft speech:

A Papa Hopper [əˌpʰɑpəˈhɑpɚ],

A Papa Hippie [əˌpʰɑpəˈhɪpi],

A Papa Hooper [əˌpʰɑpəˈhupɚ].

The focus of analysis was the intervocalic /h/ initiating the fourth, primary stressed syllable. The target /h/ was placed at the beginning of a stressed syllable to decrease the likelihood that it would be lenited or deleted (Pierrehumbert and Talkin, 1992). The differing vowel contexts were intended to induce variation in supraglottal resistance. Small variation in $f0$ might also occur due to intrinsic $f0$ effects (e.g., Whalen and Levitt, 1995). The loudness conditions were intended to yield a range of subglottal pressures for each speaker.

Participants were asked to use their typical conversational speech rate throughout the experiment. During recording, the investigator or an assistant presented each utterance orally, and the speaker then repeated it five times. Each utterance appeared five times per loudness block, with utterances randomized within block. Thus, for each speaker, approximately 225 tokens of /h/ were collected: Five repetitions per 45 input trials (3 utterances × 3 loudness conditions × 5 presentations per loudness block). The first block of utterances was produced at normal loudness, described as “What you would use for normal conversation.” The second block was produced in loud voice, described as “What you would use for a person in the next room.” To encourage louder speech, the investigator or assistant stood several feet away from the speaker throughout this condition and provided the utterances in a loud voice. The soft condition was recorded last. Here, speakers were asked to talk in a voice that they would use with “a person sitting with his/her ear very close to your mouth.” Subjects were explicitly asked not to whisper in this condition. To elicit soft speech, the investigator/assistant sat close to the subject and presented the utterances in a soft voice.

For one participant (F1), a fourth block was recorded, consisting of a longer carrier phrase (“Mama Papa Hopper/Hippie/Hooper”) produced at normal loudness. The purpose of this manipulation was to determine whether /h/ voicing varied as a function of position in utterance. Although this speaker demonstrated lower subglottal pressures in this block than in her first block at normal loudness, there were no significant differences in the voice timing measures, so the two blocks were combined for the analyses reported here.

C. Equipment and recording

Three signals were recorded for each utterance. An acoustic signal was recorded using a directional microphone (Sennheiser MKH816T) positioned approximately 2 feet from the speaker. The acoustic signals were low-pass filtered at 9.5 kHz and sampled at 20 kHz. Two aerodynamic signals were collected, filtered at 4.8 kHz, and sampled at 10 kHz. Oral airflow was collected using a Rothenberg mask and Glottal Enterprises hardware (MSIF-2). Speakers were reminded throughout recording to keep the mask pressed firmly to the face to prevent leaks, and the investigator or assistant visually assessed mask fit throughout recording. An

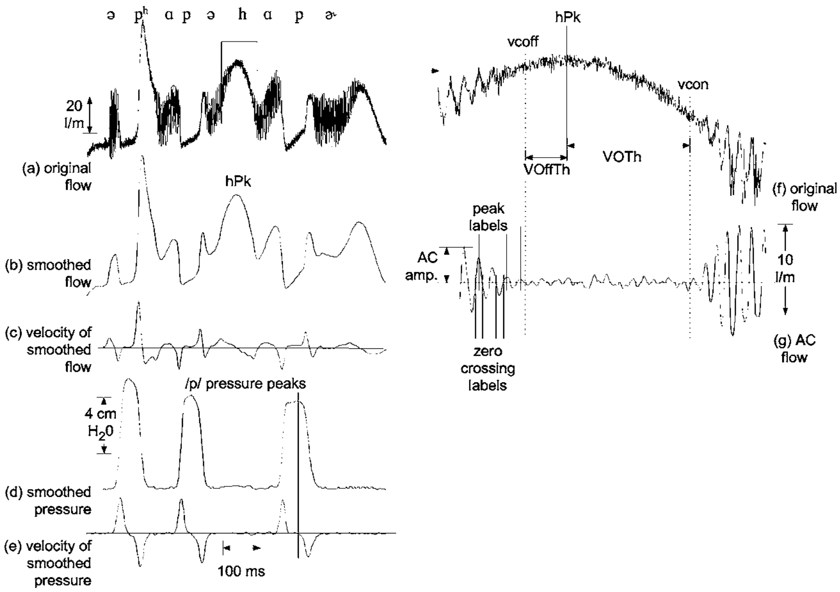

FIG. 1. One token from speaker F1, showing events used in measurement. The token shown here has a voicing break of approximately 70 ms. Panel (a) Original flow signal (lightly smoothed to remove noise). Panel (b) Smoothed or DC flow signal. Panel (c) First derivative (velocity) of the smoothed flow. Panel (d) Smoothed pressure signal. Panel (e) First derivative (velocity) of smoothed pressure. Panel (f) A short segment of the original flow signal expanded to show voicing offset and onset. Panel (g) The AC flow signal for the same time frame shown in panel (f).

intraoral pressure signal was obtained using a catheter-tip pressure transducer (Gaeltek CT/S) positioned within a piece of medical tubing fed through one of the holes in the mask's front. A plastic clamp was screwed tightly to the distal end of the tube, securing the transducer inside the tube and ensuring an airtight seal. Tube length and angle were adjusted at the beginning of the experiment so that the tube was positioned just inside the oral cavity, past the teeth, but not so far as to interfere with articulation. Occasionally, a speaker noted that, even after adjustment, she could still feel the tube during production of the high vowels, especially /ɪ/. In these cases, tube position was considered acceptable when (a) the speaker was not uncomfortable; (b) she felt that the tube did not alter her speech patterns; and (c) the pressure signal showed expected shape variations during oral closures.

Calibration signals were obtained immediately after each input session using a rotameter for the flow and a standing-water manometer for the pressure. Straight-line fits were obtained from these signals and the slope coefficient was applied to the raw signals as the first step in data processing. To correct for low-level drift in the signals over the course of the recording session, a unique baseline (intercept) was obtained for each pressure and flow signal (five repetitions in response to one utterance presentation). The /p/ closures (during which there is no airflow) were used to set the zero flow level for each signal, whereas the stressed vowels (when the vocal tract is most open) were used to set the zero-pressure level.

D. Signal processing

Signals were transferred to a VAX computer for analysis. The acoustic signal was used only for auditory assessment of the productions; all measurements were performed on the airflow, air pressure, or signals derived from them. Processing involved the following steps:

(a) The calibrated flow signals were lightly smoothed with a 5-point triangular window (0.5 ms) to remove low-amplitude, high-frequency noise in the signals. Henceforth, we will refer to these lightly smoothed signals as the "original flow signals."

(b) Smoothed versions of the flow and pressure were obtained by smoothing the signals twice with a 133-point triangular window. This process obliterated all or most evidence of glottal pulses. The smoothed flow signal shows low-frequency variation over the course of the utterance, reflecting articulatory movements, including vocal-fold abduction. For ease of reference, we will also refer to this smoothed, slowly varying signal as the DC flow, to contrast it with the rapidly varying AC flow signal that reflects vocal-fold vibration.

(c) AC flow signals were generated by subtracting the smoothed flow signal from the original (lightly smoothed) flow signal, and smoothing the result with a 25-point triangular window.

(d) First time derivatives (velocity signals) were estimated from the smoothed flow and pressure signals using a 3-point difference algorithm, and smoothed iteratively (typically twice) with a 133-point triangular window until zero crossings for major aerodynamic events could be easily obtained.

E. Measurements

The following measures were made for each token (see Fig. 1):

(a) Peak airflow during abduction for /h/ (hPk) was defined by the zero crossing in the first derivative of the smoothed flow signal [Fig. 1, panels (b) and (c)]. In the simple case of nonrotational flows through an open vocal tract, the baseline or DC flow signal at the mouth provides a good approximation to the low-frequency flow at glottal exit, which is in turn proportional to glottal area when other factors (chiefly, air density and subglottal pressure) are held constant (Titze, 1988). Thus, the baseline flow variation during /h/ reflects changes in glottal cross-sectional area, and the hPk measure pro-

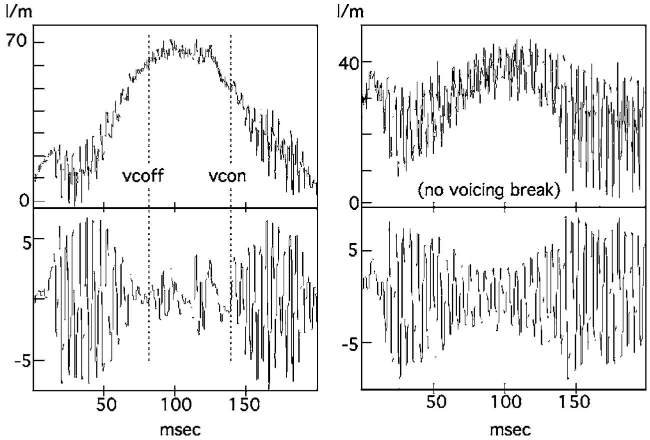

FIG. 2. Two signals from speaker F2, showing the original flow signals (above) and the AC flow signals (below). The token at left shows irregular flow variations around the /h/ flow peak. The voicing offset (vcoff) and onset (vcon) labels mark the region considered devoiced in this token. The token at right shows an /h/ with regular voicing continuing throughout.

vides an estimate of maximum glottal abduction. We discuss factors that affect the accuracy of this estimate below in Sec. II F.

(b) Peak pressures were measured during the second and third /p/ closures (i.e., those preceding and following the /h/). As with hPk, these were obtained via zero crossings in the first derivative of the smoothed pressure signal [see Fig. 1, panels (d) and (e)]. The two values were averaged to provide an estimate of subglottal pressure, or P_{sub} (cf. Löfqvist *et al.*, 1982; Smitheran and Hixon, 1981).

(c) Times of voicing offset and onset were determined by visual inspection of the AC flow signals, with reference to the original flow signals [Fig. 1, panels (f) and (g)]. Voicing was defined as periodic oscillation in the AC flow signal. Some tokens in some speakers showed AC flow variations around the /h/ flow peak that were highly irregular in frequency and amplitude. These were considered to represent unstable, chaotic movements of the vocal folds rather than the sustainable, regular vibration characteristic of voicing. An example of such chaotic oscillation is shown in Fig. 2, along with a fully voiced /h/ from the same speaker.

Two duration measures were calculated from the voicing offset and onset times: Voicing offset to hPk (VOffTh) and hPk to voicing onset (VOTh). VOTh is analogous to the voice onset time (VOT; Lisker and Abramson, 1964) in an aspirated stop consonant, in which the peak abduction occurs at approximately the time of oral release (cf. Löfqvist, 1992 for some qualifications). For tokens of /h/ with a voicing break, VOTh and/or VOffTh were positive. For cases of fully voiced /h/, both VOTh and VOffTh were approximately 0; specifically, the voicing offset and onset labels were placed within one glottal pulse of the hPk, as described below.

Presence or absence of a voicing break was also determined based on the times of voicing offset and onset. The automatic measurement routines for $f0$ and pulse (AC) amplitude required a whole number of glottal pulses before and after voicing offset/onset, or, in cases of fully voiced /h/, the time of peak abduction (hPk). Glottal pulses whose open phase coincided with the hPk label were thus not measured. Such tokens had a measured "voicing break" duration of about a pulse period. To prevent counting these tokens as instances of devoiced /h/, a two-pulse-period criterion was defined for each subject, based on the session average of her $f0$ values at voicing offset and onset (an average of three pulses, as described below). Tokens with a voicing break two periods or longer were labeled as "devoiced."

(d) Flow amplitudes in the smoothed (DC) signal were measured at the times of voicing offset and onset (DCOff, DCOn). These are the amplitude values that correspond to the voicing offset and onset times indicated in Fig. 1, panel (f).

(e) Fundamental frequency ($f0$) at the times of voicing offset and onset ($f0$Off, $f0$On) was obtained from zero crossings in the phonated regions of the AC flow signal [Fig. 1, panel (g)]. Two measures of $f0$ were obtained at both offset and onset: The $f0$ of the last/first pulse, and an average of the last/first three pulses. As discussed below, the average measures proved to be more reliable, and were therefore used for the final statistical analyses.

(f) Glottal pulse (AC flow) amplitudes immediately before/after voicing offset/onset (ACOff, ACOn) were obtained by performing peak picking in the AC flow signal [Fig. 1, panel (g)]. As with $f0$, two measures were obtained at offset and onset: The amplitude of the last/first pulse, and an average of the last/first three pulses; the statistical analyses ultimately included only the average measures.

In summary, the measures provide information on (a) the degree of abduction during /h/ (hPk); (b) the subglottal pressure level during the utterance (P_{sub}); (c) voice timing (VOffTh, VOTh); (d) the degree of abduction at which voicing stopped and started (DCOff, DCOn); (e) changes in passive and actively induced longitudinal tension within the folds, as reflected by $f0$ measures ($f0$Off, $f0$On); and (f) the vibratory amplitude of the vocal folds (ACOff, ACOn).

F. Accuracy of estimated measures

As noted in the previous section, we use the hPk measure as an estimate of the extent of glottal abduction during /h/, and the average of the intraoral pressure peaks for the two /p/'s flanking the VhV sequence as an estimate of subglottal pressure (P_{sub}). Two factors limit the precision of these estimates.

In the case of hPk, the qualification involves the extent to which P_{sub} remains constant during abduction. When P_{sub} is stable, airflow varies in proportion to the extent of glottal opening, and variations in hPk directly reflect changes in glottal opening. In fact, P_{sub} may decrease somewhat when the glottis opens and the vocal tract is open (Löfqvist, 1975; Ohala, 1990); as a result, the airflow changes may underestimate glottal area change. Few direct measurements exist to quantify the extent of such a decrease across speakers, but in our recent modeling work (Lucero and Koenig, 2005) we used P_{sub} as a control parameter for reproducing voicing patterns in intervocalic /ɑhɑ/ produced by eight speakers (including some of the women analyzed here). The results of those simulations showed an average P_{sub} decrease across speakers of about 2 cm H_20 during the /h/ compared to the neighboring vowel. The corresponding value obtained by Löfqvist (1975) was about 1.2 cm H_20.

In the case of P_{sub}, it should be noted that the /h/ in our utterances introduced a stressed syllable, whereas the pressure peaks were measured during /p/ closures initiating unstressed syllables. One effect of stress is to increase subglottal pressure momentarily. Past studies on subglottal pressure variation as a function of stress suggest that stressed syllables are typically produced at pressures 1–2 cm H_20 higher than unstressed syllables (Brown and McGlone, 1974; Dixit and Brown, 1978; Dixit and Shipp, 1985; McGlone and Shipp, 1972).

Decreases of P_{sub} as a function of abduction and increases as a function of stress are thus similar in magnitude. In combination, therefore, the two factors should roughly balance out in our data. We note two additional points about these measures. In the case of hPk, the effect of a subglottal pressure decrease during abduction should become more extreme as abduction extent becomes larger. The implication is that the hPk measure should be considered a conservative measure of abduction degree; in other words, the hPk variation in our data is probably more limited than the actual variation in abduction degree. As for P_{sub}, our intent was to capture major variations in pressure as a function of loudness condition, and the statistical analyses generally show the expected pressure variation across loudness conditions.

G. Statistics

Paired *t*-tests were used to test for offset–onset differences. To explore the effects of vowel and loudness condition on measured variables, analyses of variance (ANOVAs) and χ^2 analyses were performed. Finally, simple correlations, multiple regression, and principal components analyses were performed for each subject in order to clarify the relationships among correlated variables. For these relational analyses, the loudness condition was recoded into a quasicontinuous variable with loud=1, normal=0, soft=−1; the vowel variable was recoded into two variables, one with /ɑ/=1, /u/=/ɪ/=0 ("/ɑ/-ness") and one with /ɪ/=1, /ɑ/=/u/=0 ("/ɪ/-ness"). An alpha level of 0.01 was set as the significance criterion for all analyses.

H. Measurement reliability

A subset of the data was remeasured by the same investigator (the first author) several weeks after the original measurements were completed (for some subjects, a duration of several months). Specifically, voicing offsets and onsets, which were determined visually, were remeasured for the third repetition (out of five) in the even-numbered input trials (22 of 45) for each speaker. The set of remeasured data comprised 128 tokens, or approximately 9% of the data. All measures associated with voicing offsets and onsets (DC and AC flow, *f*0) were also derived again. Pressure peaks during /p/ and flow peaks during /h/ were not remeasured because they were defined automatically rather than visually. The original and remeasured values were compared by performing a correlation on the two sets of data, and by calculating means and standard deviations of the differences between each pair of measurements.

TABLE II. Results of reliability analyses: *r*- and *p* values from the correlation analysis, and the average and standard deviation (SD) of the differences between the original and remeasured data sets. Durational measures (VOffTh, VOTh) are in ms; flow values (DCOff/On, ACOff/On) are in 1/m; *f*0 values are in Hz. VOffTh=time of voicing offset, relative to the peak flow in /h/; VOTh=time of voicing onset relative to the peak flow in /h/; DCOff and DCOn refer to the baseline, or DC, flow values at voicing offset and onset; ACOff and ACOn refer to the pulse amplitudes at offset and onset; *f*0Off and *f*0On refer to the *f*0 values at offset and onset.

Variable	*r*	*p*	Mean diff.	SD of diff.
VOffTh	0.95	<0.001	−0.56	4.61
VOTh	0.95	<0.001	0.12	6.30
DCOff	1.00	<0.001	0.03	2.37
DCOn	0.99	<0.001	<0.01	3.22
ACOff	0.91	<0.001	0.10	0.85
ACOn	0.92	<0.001	0.08	1.01
*f*0Off	0.91	<0.001	−0.42	11.05
*f*0On	0.82	<0.001	−0.51	19.37

The correlations between the original and the remeasured data were highly significant, with $p<0.001$ for all measures, but some differences were observed between the single and the three-pulse averaged measures performed for AC flow and *f*0. In the case of AC flow, the two measurement sets were similar, with all *r* values between 0.8 and 0.92. However, the *r* values for the single-pulse *f*0 measures were considerably lower than those for the three-pulse averaged measures: Single-pulse *r* values were 0.56 for offsets and 0.69 for onsets, but 0.91 for offsets and 0.82 for onsets in the averaged measures. In order to maintain high reliability across measures, we opted to use the averaged measures in all subsequent analyses. The reliability results for the final set of measures are given in Table II.

Because laryngeal conditions change rapidly in the vicinity of an abduction gesture, the three-pulse average measures do represent some loss of information on conditions around voicing thresholds. To determine how closely the averaged measures captured the patterns of the single-pulse measures, we also computed correlations between the single-pulse and averaged measures for the entire dataset (all productions, all speakers). All correlations were significant at $p<0.001$; *r* values ranged from 0.8–0.87. When the correlations were run within single speakers, the *r* values varied from 0.53–0.92, but all were again highly significant at $p<0.001$.

III. RESULTS

A. Offset–onset differences

Past work has indicated that vocal-fold vibration is subject to a hysteresis effect, whereby voicing offsets and onsets occur under different conditions. To investigate offset–onset differences in the current data, paired *t*-tests were performed

TABLE III. Offset–onset differences in tokens with voicing breaks for all speakers F1–F6: Means, SDs, and results of two-tailed t-tests. Temporal measures (VOffTh, VOTh) are in ms; amplitude (DC, AC) measures are in 1/m; $f0$ is in Hz.

	F1	F2	F3	F4	F5	F6
VOffTh means (SD)	9.3 (14.3)	6.5 (13.3)	8.7 (13.0)	4.0 (7.2)	15.2 (14.4)	22.0 (13.9)
VOTh means (SD)	45.9 (18.8)	33.1 (18.1)	27.5 (16.7)	24.5 (10.0)	39.9 (18.8)	47.0 (19.2)
t	−18.64	−9.83	−5.59	−21.04	−11.03	−12.62
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
DCOff means (SD)	60.4 (12.9)	51.6 (19.0)	15.5 (5.2)	35.0 (16.8)	17.5 (7.5)	37.3 (13.6)
DCOn means (SD)	43.2 (13.4)	42.3 (20.3)	13.6 (5.7)	23.5 (16.1)	11.6 (7.9)	27.6 (11.9)
t	20.85	8.48	4.04	15.30	14.15	13.40
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
$f0$Off means (SD)	175.2 (12.0)	195.7 (17.1)	184.0 (15.9)	209.4 (14.4)	135.8 (19.4)	208.3 (16.4)
$f0$On means (SD)	204.3 (14.5)	216.0 (15.7)	205.1 (15.1)	243.5 (16.6)	179.5 (25.7)	257.8 (32.1)
t	−20.82	−10.25	−9.02	−20.18	−18.61	−18.44
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
ACOff means (SD)	1.8 (0.9)	3.4 (2.0)	0.9 (0.6)	3.2 (1.9)	2.6 (1.1)	1.3 (1.1)
ACOn means (SD)	2.9 (1.6)	3.6 (1.9)	1.1 (0.7)	4.3 (2.8)	2.3 (1.1)	1.8 (1.6)
t	−8.33	−0.93	−2.56	−5.6	2.08	−5.63
p	<0.001	0.357	0.013	<0.001	0.040	<0.001

on the following sets of measures in tokens with voicing breaks: VOffTh vs VOTh; DCOff vs DCOn; $f0$Off vs $f0$On; and ACOff vs ACOn. Results, given in Table III, indicate that, for all speakers (a) voicing offsets occurred significantly closer in time relative to the flow peak than voicing onsets (VOffTh < VOTh); (b) voicing offsets occurred at significantly higher DC flow amplitudes than voicing onsets (DCOff > DCOn); and (c) voicing began at significantly higher levels of $f0$ than it ceased ($f0$Off < $f0$Off). Qualitatively, five of the six speakers had lower AC flows at voicing offset than onset, but the difference was only significant in three speakers. It may be that the use of the three-pulse average measure masked short-term changes in pulse amplitude near voicing offset and onsets. The results for voice timing and DC flow amplitudes at voicing offset and onset are consistent with our expectations of a hysteresis effect. In these data, the higher $f0$ following /h/ most likely reflects the stress pattern of the utterance, namely the fact the target /h/ initiated a stressed syllable.

A final observation from Table III is that standard deviations (SDs) tended to be higher overall for voicing onsets than for offsets. The t-tests on the 6 SDs for each measurement pair showed that the difference was significant for VOffTh vs VOTh [$t(5)=11.795$, $p<0.0001$; mean difference=4.25]. Given the small number of data points for each comparison, however, the results of these analyses must be considered preliminary.

B. Categorical effects of loudness and vowel on devoicing

Figure 3 shows the percentage of devoiced /h/ for each speaker as a combined function of vowel and loudness. The statistical tests for loudness- and vowel-related changes are summarized in Table IV. Full results of the ANOVAs, with means and standard deviations for the measures, are given in the Appendixes. In the following section, we present an overview of the results; individual patterns are considered in more detail below.

Figure 3 and Table IV indicate that the vowel and loud-

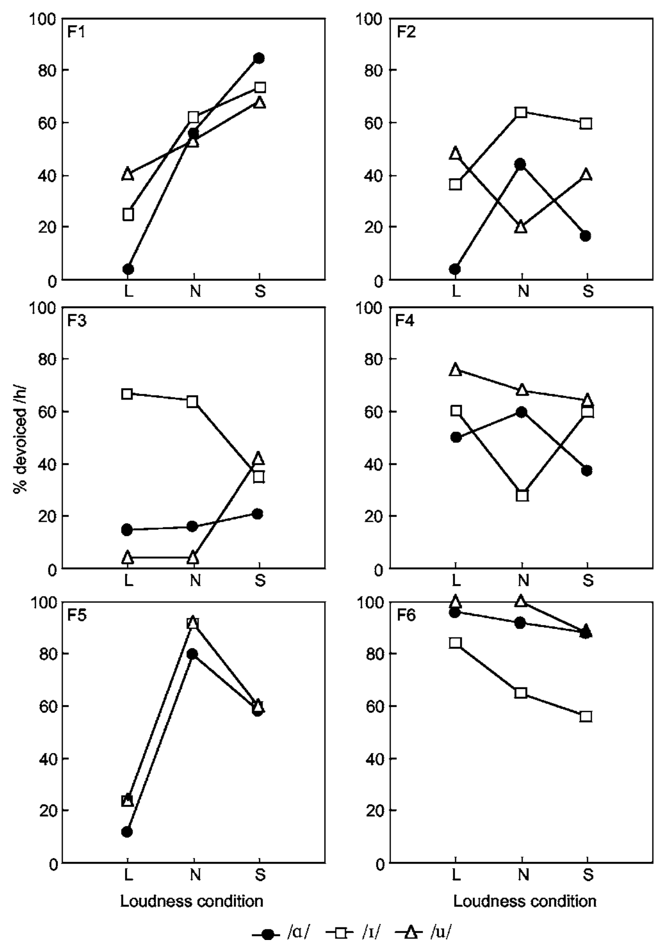

FIG. 3. Percentage of voiceless /h/ as a function of loudness condition and vowel for each speaker. L=loud; N=normal; S=soft.

TABLE IV. Summary of vowel and loudness effects. Percentage of devoicing across conditions was assessed using χ^2 analyses; effects of loudness and vowel on P_{sub} and hPk measures were assessed using ANOVAs. Asterisks indicate significance at $p<0.01$. The results of the post-hoc tests are summarized in cases of significant p-values. For example, speaker F1 had significant P_{sub} differences among all 3 of her loudness conditions; for F2, loud and normal were not significantly different from each other, but both were significantly different from the soft condition.

Speaker	% devoicing as a function of loudness	% devoicing as a function of vowel	P_{sub} as a function of loudness condition	hPk as a function of loudness condition	P_{sub} as a function of vowel	P_{sub} vowel-by-loudness interaction
F1	*		* $L>N>S$			
F2		*	* $L,N>S$	* $L>N,S$	* /ɑ,u/ > /ɪ/	
F3		*	* $L>N,S$		* /u/ > /ɑ,ɪ/	($p=0.012$) Average pressure pattern /u/ > /ɑ,ɪ/ true of loud and normal; in the soft condition, /u/ had the lowest pressure, and /ɪ/ had the highest.
F4		($p=0.02$)	* $L>N>S$	* $S>N,L$		* Steeper pressure decrease for /ɪ/ from loud to normal than for /ɑ, u/
F5	*		* $L,S>N$	* $L>N,S$	* /ɪ,u/ > /ɑ/	* $L>S>N$ differences smaller for /u/
F6	($p=0.02$)	*	* $L,N>S$	* $L>N>S$		* $L>N>S$ for /ɪ/ and /u/; for /ɑ/, ($L=N>S$).

ness conditions had different effects on devoicing across speakers. To understand these effects, we must consider how the vowel and loudness manipulations affected underlying production parameters. The loudness manipulation was chiefly intended to yield variation in subglottal pressure (P_{sub}). To determine whether this was achieved, ANOVAs were run on P_{sub} as a function of loudness level (L=loud, N=normal, S=soft). The results (summarized in Table IV) showed that P_{sub} varied significantly ($p<0.001$) across loudness condition for all speakers. Five of the six speakers presented the expected pattern of $L>N>S$, although not all pairwise tests were significant. In one speaker (F5), average P_{sub}'s were higher in both the loud and soft conditions than in normal loudness ($L,S>N$). This pattern is discussed further in Sec. III D below.

Dromey and Ramig (1998) have demonstrated that speakers may change articulatory patterns as well as subglottal pressure in varying loudness conditions. To investigate whether the speakers recorded here varied their degree of abduction as loudness changed, ANOVAs were run on hPk measures as a function of loudness condition. The results (see Table IV) indicated that loudness had significant effects on hPk for four speakers, but the direction of the effect was inconsistent: One speaker (F6) showed decreasing hPk flows as loudness decreased; two (F2, F5) showed higher hPk flows for the loud condition only; and one (F4) had the highest hPk values in the soft condition.

To determine whether speakers altered P_{sub} in response to changing vocal-tract (vowel) postures, ANOVAs were run on P_{sub} as a function of vowel. The results (see Table IV) were significant for three speakers (F2, F3, F5), but the direction of the effect varied across speakers. Further, P_{sub} showed significant interactions between loudness and vowel for three speakers (F4, F5, F6; results are again summarized in Table IV). We will return to these data below, and integrate them with the results of the relational analyses.

C. Predicting VOffTh and VOTh from measured variables

1. Correlations

As indicated above, interrelationships among many variables were expected. For example, loudness condition was expected to affect P_{sub}, but might also affect $f0$, via changes in P_{sub} and/or laryngeal setting. To quantify the relationships among the experimental variables, and to determine which variables were most strongly correlated with the voice timing variables (VOffTh, VOTh), simple regression analyses were performed. The occurrence of both voiced and voiceless /h/ meant that there were two populations of data points for both VOTh and VOffTh (zero vs a range of positive values); thus, two separate analyses were performed: One including the full data set, and another including the voiceless tokens only. Comparison of the results with and without fully voiced /h/ suggested that the pattern of results was similar across the two analyses. To formally quantify the relationship between the two datasets, we performed an r-to-z transform, correcting for non-normality of r-value distributions, and ran a correlation on the data from the full and the voiceless-only analyses. Results indicated high correlations for all data analyzed together ($r=0.900$, $p<0.0001$) and for offsets and onsets analyzed separately (offsets, $r=0.918$, $p<0.0001$; onsets, $r=0.879$, $p<0.0001$). These high correlations indicate that the variables associated with voicing vs devoicing also tend to predict the duration of devoicing when it occurs.

Table V presents the correlation matrices for the full data set (including both voiced and voiceless /h/). These data reveal some consistencies across speakers, and provide fur-

TABLE V. Correlation matrices for all tokens (voiced and voiceless) in all subjects. For the sake of compactness, the left column serves both offset and onset matrices; the offset variables (VOffTh, DCOff, $f0$Off, ACOff) refer to the left matrix, and the onset variables refer to the right matrix. Ellipses indicate correlations that were not significant at $p<0.01$.

	Voicing offset									Voicing onset								
	VOffTh	/ɑ/	/ɪ/	Ld	DCOff	hPk	Pres	$f0$Off	ACOff	VOTh	/ɑ/	/ɪ/	Ld	DCOn	hPk	Pres	$f0$On	ACOn
									F1									
VOffTh/VOTh	1.00	···	···	···	0.15	0.26	···	···	···	1.00	···	···	−0.43	−0.52	0.21	−0.16	0.72	0.27
/ɑ/-ness		1.00	−0.49	···	−0.31	−0.31	···	···	···		1.00	−0.49	···	−0.22	−0.31	···	···	−0.15
/ɪ/-ness			1.00	···	0.26	0.27	···	···	···			1.00	···	···	0.27	···	···	0.27
Loudness				1.00	···	···	0.59	···	0.19				1.00	0.23	···	0.59	−0.28	···
DCOff/On					1.00	0.99	···	···	···					1.00	0.63	···	−0.33	−0.16
hPk						1.00	···	···	···						1.00	···	0.30	0.15
Pres							1.00	−0.18	···							1.00	···	···
f_0Off/On								1.00	···								1.00	0.30
ACOff/On									1.00									1.00
									F2									
VOffTh/VOTh	1.00	···	0.25	···	0.24	0.32	···	···	···	1.00	−0.24	···	···	−0.28	···	···	0.39	···
/ɑ/-ness		1.00	−0.50	···	···	···	···	···	−0.25		1.00	−0.5	···	···	···	···	−0.20	···
/ɪ/-ness			1.00	···	0.42	0.42	−0.21	···	···			1.00	···	0.34	0.42	−0.21	···	0.20
Loudness				1.00	0.52	0.50	0.22	0.32	0.42				1.00	0.56	0.50	0.22	0.25	0.40
DCOff/On					1.00	0.98	···	0.19	0.39					1.00	0.89	···	···	0.38
hPk						1.00	···	0.19	0.37						1.00	···	0.22	0.37
Pres							1.00	···	0.23							1.00	···	0.21
f_0Off/On								1.00	0.25								1.00	0.20
ACOff/On									1.00									1.00
									F3									
VOffTh/VOTh	1.00	···	0.45	···	···	0.27	···	···	···	1.00	···	0.19	···	···	0.20	···	0.28	···
/ɑ/-ness		1.00	−0.47	···	−0.60	−0.61	···	−0.27	−0.37		1.00	−0.47	···	−0.58	−0.61	···	−0.29	−0.25
/ɪ/-ness			1.00	···	0.44	0.53	···	···	0.41			1.00	···	0.46	0.53	···	0.28	0.54
Loudness				1.00	···	···	0.75	0.48	···				1.00	···	···	0.75	0.49	···
DCOff/On					1.00	0.95	···	0.24	0.60					1.00	0.91	···	···	0.48
hPk						1.00	···	0.23	0.61						1.00	···	0.27	0.54
Pres							1.00	0.67	···							1.00	0.57	···
f_0Off/On								1.00	0.29								1.00	0.32
ACOff/On									1.00									1.00
									F4									
VOffTh/VOTh	1.00	···	···	···	···	···	···	···	···	1.00	···	···	···	−0.48	···	···	0.68	0.26
/ɑ/-ness		1.00	−0.49	···	···	···	···	0.21	−0.16		1.00	−0.49	···	···	···	···	···	···
/ɪ/-ness			1.00	···	···	···	···	···	0.10			1.00	···	···	···	···	···	···
Loudness				1.00	−0.70	−0.70	0.93	···	0.28				1.00	−0.61	−0.70	0.93	0.34	0.25
DCOff/On					1.00	0.99	−0.61	···	−0.30					1.00	0.85	−0.53	−0.43	−0.44
hPk						1.00	−0.60	···	−0.31						1.00	−0.60	···	−0.31
Pres							1.00	0.22	0.23							1.00	0.41	0.22
f_0Off/On								1.00	−0.36								1.00	···
ACOff/On									1.00									1.00
									F5									
VOffTh/VOTh	1.00	···	···	···	−0.27	···	−0.22	···	···	1.00	···	···	−0.27	−0.64	···	−0.25	0.31	−0.21
/ɑ/-ness		1.00	−0.49	···	−0.47	−0.49	−0.29	···	−0.35		1.00	−0.49	···	−0.30	−0.49	−0.29	···	−0.30
/ɪ/-ness			1.00	···	0.43	0.50	···	0.19	0.32			1.00	···	0.36	0.50	···	···	0.23
Loudness				1.00	0.30	0.27	···	0.72	0.40				1.00	0.39	0.27	···	0.58	0.55
DCOff/On					1.00	0.95	0.30	0.26	0.45					1.00	0.76	0.35	···	0.41
hPk						1.00	0.22	0.23	0.44						1.00	0.22	0.25	0.42
Pres							1.00	0.19	0.25							1.00	···	0.27
f_0Off/On								1.00	0.70								1.00	0.58
ACOff/On									1.00									1.00

TABLE V. (*Continued.*)

	Voicing offset									Voicing onset								
	VOffTh	/ɑ/	/ɪ/	Ld	DCOff	hPk	Pres	$f0$Off	ACOff	VOTh	/ɑ/	/ɪ/	Ld	DCOn	hPk	Pres	$f0$On	ACOn
									F6									
VOffTh/VOTh	1.00	0.35	⋯	0.20	⋯	0.19	0.21	⋯	⋯	1.00	⋯	−0.47	⋯	−0.48	⋯	⋯	⋯	⋯
/ɑ/-ness		1.00	−0.49	⋯	−0.42	−0.29	⋯	⋯	⋯		1.00	−0.49	⋯	−0.21	−0.29	⋯	⋯	⋯
/ɪ/-ness			1.00	⋯	0.37	0.36	⋯	⋯	⋯			1.00	⋯	0.65	0.36	⋯	⋯	⋯
Loudness				1.00	0.51	0.59	0.51	−0.24	⋯				1.00	0.42	0.59	0.51	⋯	⋯
DCOff/On					1.00	0.93	0.44	−0.30	⋯					1.00	0.74	0.36	0.18	⋯
hPk						1.00	0.48	−0.28	⋯						1.00	0.48	⋯	⋯
Pres							1.00	−0.18	0.23							1.00	0.38	0.33
f_0Off/On								1.00	⋯								1.00	0.44
ACOff/On									1.00									1.00

ther evidence for offset–onset differences. First, in five of six speakers (all but F3), VOTh had a significant negative correlation with DCOn. This reflects the fact that a longer VOTh corresponds to voicing occurring later in the adduction gesture, and at lower flow rates. Correlations between VOffTh and DCOff were weaker and significant for only three of six speakers. However, DCOff and hPk were highly, significantly correlated for all speakers: For both the full and the voiceless-only analysis, all r values were above 0.89. (Recall that in tokens with no voicing break the times of voicing "offset" and "onset" were set within a pulse of the /h/ flow peak, so that offset, onset, and peak values were virtually identical; for this analysis, therefore, the voiceless-only dataset is more valid.) These high correlations were also expected given that voicing offset typically occurred very close in time to the peak abduction in /h/. The comparable onset correlations (DCOn−hPk) were significant for all speakers, but the r values were lower than for offsets (0.63–0.89 in the voiceless-only analysis).

Apart from these consistencies, the variables correlated with VOffTh and VOTh to different degrees across speakers. For example, VOTh in speaker F4 correlated with DCOn, $f0$On, and, to a lesser degree, ACOn. For F5, VOTh correlated with loudness, DCOn, pressure, $f0$On, and ACOn; note that the direction of the AC effect is the opposite of that for F4. Speaker F6 showed no effect of loudness, pressure, $f0$On, or ACOn; instead, VOTh values correlated with the /ɪ/ context along with DCOn.

2. Multiple regression

To determine how strongly each of the experimental variables predicted voice timing in each speaker, within-subject multiple regression (MR) analyses were performed, with dependent variables of VOffTh and VOTh. Table VI shows the results for the full data set (voiced and voiceless tokens). As with the correlational analyses, results were similar for the full dataset as compared to the voiceless-only analysis, although the F values were lower for the smaller (voiceless only) data set. These analyses supplement the principal components analysis (discussed below), which is not inferential and does not provide statistical significance levels.

Table VI indicates that the measured variables significantly predicted the times of voicing offset and onset in all speakers, but the r values and F values were in all cases higher for voicing onset than voicing offset. The t-tests on the log-transformed r values indicated that these offset–onset

TABLE VI. Results of multiple regression analyses for the full data set (voiced and voiceless tokens together), with VOffTh and VOTh as dependent variables. Double asterisks (⋆⋆) indicate cases of $p < 0.001$; single asterisks (⋆) indicate $p < 0.01$; tildes (∼) indicate $p < 0.05$.

	F1		F2		F3		F4		F5		F6	
	Off	On	Off	On	Off	On	Off	On	Off	On	Off	On
Overall r	0.66	0.89	0.58	0.9	0.74	0.91	0.67	0.94	0.69	0.9	0.74	0.86
Overall F	28.49	144.57	13.36	110.48	29.83	110.37	21.49	194.59	24.02	111.86	31.23	72.14
p value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Intercept								⋆⋆				⋆⋆
/ɑ/-ness			⋆⋆	∼	∼	∼	⋆	∼				⋆⋆
/ɪ/-ness		⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆					⋆	∼
Loud		⋆⋆						⋆⋆				
DCOff/On	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆
hPk	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆
Pres							⋆⋆					
f_0Off/On	∼	⋆⋆		∼				⋆⋆		⋆⋆		
AC3Off/On					∼					⋆⋆		

differences were significant [$t(5)=6.836$, $p=0.001$; mean difference=0.283]. Again, given the small number of data points for these tests, the conclusions must be considered tentative. The multiple regression also shows that the contributions of DCOff/DCOn and hPk were significant in all cases. The universal significance of DCOff and DCOn as predictors of VOffTh and VOTh (respectively) implies that, as one would expect, the degree of abduction is a major factor in determining voice timing. Otherwise, the results are mixed, suggesting that other variables contribute to voicing control to differing degrees across subjects.

3. Principal components analysis (PCA)

The general goal of principal components analysis is to determine the underlying dimensionality among a large number of correlated variables. Loadings on a single factor represent variables that are intercorrelated; conversely, correlations are minimized across factors (Dunteman, 1989). Thus, the PCA results provide greater insight into how variables associated with voicing control are interrelated within a speaker.

Table VII presents the orthogonal solutions yielded by principal components analysis, showing the results for all data (voiced and voiceless tokens). The factors were rotated using the Varimax procedure. The number of factors retained for each analysis was determined using a composite procedure, in order to minimize dependence upon a single method and criterion. Specifically, factors were retained depending on the larger of two values as determined by (a) the change point on a scree plot (Cattell, 1966) and (b) a 75% total variance criterion (cf. Dunteman, 1989; Jackson, 1991).

The factor solutions effectively characterized the data sets for all subjects (all χ^2 significant at $p<0.001$). This was true when the analysis was run on the entire data set and when voiced tokens were removed, though the χ^2's were lower in the latter case. In both the full and voiceless-only analyses, four factors were extracted for all subjects except F3, for whom three factors characterized her dataset. Table VII indicates that certain factors group together across subjects: In particular, all subjects have at least one factor with moderate to heavy loadings of hPk and DCOff/On, for both voicing offsets and onsets. Other variables loading on these factors differ across speakers. For example, speaker F1 (factor 1, voicing offsets) has DCOff and hPk combined with AC flow only, whereas in F4 (factor 1, voicing offsets) this factor includes loudness and pressure. Cross-speaker differences are also observed in the variables that load together with the voice timing variables (VOffTh, VOTh), in the number of factors the voice timing variables are loaded on, and in the factor groupings for voicing offsets vs onsets. For speaker F2, for example VOffTh loads with vowel quality (/ɪ/-ness), DCOff, and hPk, whereas for speaker F3, the factor that includes VOffTh loads only on vowel quality (/ɪ/-ness), with no other contributing factors.

Overall, the results of the relational analyses reveal a few consistencies across speakers in the variables associated with voice timing; specifically, voice timing was found to be related to the DC flow variables in all speakers, and the DC flow variables were highly correlated with each other (i.e., they loaded on common factors in the PCA). At the same time, the data indicate considerable interspeaker variability in the other factors expected to affect voicing thresholds. We explore this variability in more detail in the next section.

D. Individual speaker patterns

In this section, we bring together all the data and attempt a detailed explanation of voicing behavior in three of the six speakers (F1, F3, F5). These speakers, chosen arbitrarily from among the six, illustrate some of the ways in which speakers can differ in phonatory function.

Speaker F1 had a significant effect of loudness on devoicing, with more devoicing as loudness and P_{sub} decreased (see Fig. 3 and Table IV). The MR analysis showed vowel quality, loudness, and $f0$ to contribute to voicing onset along with DCOn and hPk. The PCA showed that loudness and pressure loaded on a single factor, with no other contributing variables. The vowel variables loaded together with AC flow, and the factors that included VOffTh and VOTh also included $f0$. Taken together, these results suggest that the loudness manipulation had the simple effect of increasing subglottal pressures in this speaker; that the vowel manipulation affected the amplitude of vocal-fold vibration; and that $f0$ made an independent contribution to voice timing.

Speaker F3 showed significant vowel effects on devoicing, and P_{sub} varied with vowel as well as loudness condition. In the loud and normal conditions, she had the least devoicing in the /u/ context (see Fig. 3), whereas in the soft condition this vowel showed the most devoicing. The devoicing pattern follows the P_{sub} data for this speaker; although her loudness-by-vowel interaction did not reach significance ($p=0.012$; cf. Table IV), qualitatively /u/ had the lowest P_{sub}'s in the soft condition, whereas it had the highest P_{sub}'s in the normal and loud conditions. Vowel effects are also evident in the MR results, where voice timing was predicted by vowel quality along with the usual hPk and DC flow. In the PCA (a) loudness, P_{sub}, and $f0$ loaded on a single factor; (b) the vowel variables loaded with AC flow as well as hPk and DC flows; and (c) the voice timing variables load with vowel (offsets) and $f0$ (onsets). The PCA loadings suggest that, for this speaker, varying vowel quality affected laryngeal abduction for /h/ (measured by DC flow) as well as on vocal-fold vibratory amplitudes (AC flow), which together yielded the significant vowel effects on phonation. The factor loadings also suggest that the loudness manipulation may have affected voicing partly via effects on $f0$.

Speaker F5 was distinguished by an unusual pattern of pressure variation in which the loud and soft conditions had higher pressures than normal loudness (Fig. 3, Table IV). Her higher amount of devoicing in the normal condition follows from the lower pressures. In the PCA, loudness is grouped together with $f0$ and AC flow, and the MR analysis shows that $f0$ and AC flow variables predicted voicing onsets along with DC flow and hPk. For this speaker, the loudness manipulation appears to have had complex effects on glottal function, affecting $f0$ and vibratory amplitudes along with simple driving pressures. In the soft condition, this speaker may have increased P_{sub} in response to greater glottal leak-

TABLE VII. Results of principal components analysis, including all tokens (voiced and voiceless) in the analysis. For clarity, only loadings of 0.4 or greater are shown.

	Voicing offset				Voicing onset			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
F1								
VOffTh/VOTh				0.76	0.91			
/ɑ/-ness			−0.74					−0.75
/ɪ/-ness			0.79					0.86
Loudness		0.89					0.86	
DCOff/On	0.93				−0.45	0.85		
hPk	0.94					0.90		
Pres		0.88					0.90	
f_0Off/On				−0.67	0.89			
ACOff/On	−0.40		0.52		0.41			0.54
F2								
VOffTh/VOTh			0.74				0.78	
/ɑ/-ness		−0.90				−0.83		
/ɪ/-ness		0.75	0.47			0.83		
Loudness	0.82				0.72			
DCOff/On	0.64		0.67		0.95			
hPk	0.62		0.71		0.89			
Pres	0.44			−0.70				0.83
f_0Off/On	0.46			0.74			0.86	
ACOff/On	0.73				0.41			0.62
F3								
VOffTh/VOTh			0.94				0.92	
/ɑ/-ness	−0.76				−0.73			
/ɪ/-ness	0.55		0.64		0.73			
Loudness		0.88				0.90		
DCOff/On	0.93				0.90			
hPk	0.91				0.93			
Pres		0.93				0.92		
f_0Off/On		0.78				0.73	0.45	
ACOff/On	0.72				0.66			
F4								
VOffTh/VOTh				0.99		0.92		
/ɑ/-ness		−0.86					−0.87	
/ɪ/-ness		0.83					0.84	
Loudness	0.93				0.96			
DCOff/On	−0.89				−0.66			0.56
hPk	−0.89				−0.79			0.43
Pres	0.88				0.94			
f_0Off/On			0.87			0.87		
ACOff/On			−0.73					−0.88
F5								
VOffTh/VOTh			0.86				−0.93	
/ɑ/-ness	−0.68			−0.47	−0.75			−0.43
/ɪ/-ness	0.79				0.86			
Loudness		0.86				0.85		
DCOff/On	0.82				0.51		0.77	
hPk	0.87				0.78			
Pres				0.91				0.93
f_0Off/On		0.94				0.89		
ACOff/On	0.40	0.70				0.76		

TABLE VII. (*Continued.*)

	Voicing offset				Voicing onset			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
				F6				
VOffTh/VOTh	0.41	0.62				−0.93		
/ɑ/-ness		0.83						−0.95
/ɪ/-ness		−0.74				0.69		0.61
Loudness	0.78				0.87			
DCOff/On	0.77	−0.47			0.62	0.70		
hPk	0.88				0.85			
Pres	0.74				0.66		0.47	
f_0Off/On				0.94			0.81	
ACOff/On			0.95				0.84	

age; visual inspection of averaged DC flow signals over the entire VCV sequence for her three loudness conditions showed higher flows during the unstressed vowel preceding the /h/ in the soft condition than for the normal or loud conditions. The soft condition also tended to have low-amplitude, sinusoidal pulse shapes, suggestive of incomplete glottal closure.

IV. DISCUSSION

A. Differences between voicing offset and onset

As expected, voicing offsets occurred closer in time to the peak abduction, and at greater degrees of abduction (higher DC flows) than voicing onsets in all speakers. This finding is consistent with a hysteresis effect, in which the requirements for initiating phonation are more stringent than those for sustaining it (e.g., Lucero, 1999; Lucero and Koenig, 2005). However, speakers varied in the extent of this effect. Across the six speakers, offset–onset differences in voice timing relative to peak abduction showed a range of 19–37 ms, with VOTh durations two to six times longer than VOffTh durations. Offset–onset differences in DC flow levels varied from 2–18 l/m; DCOff values were 1.15–1.51 times greater than DCOn values. Plant *et al.* (2004), using tracheal puncture to assess subglottal pressures at phonation offset and onset, have also reported variability in the degree of hysteresis across speakers; in their study, some subjects did not demonstrate hysteresis at all.

Several factors may contribute to this variation. One possibility is that speakers differ in anatomical parameters such as tissue damping or vocal-fold thickness. Other contributors may be the degree to which the speakers vary their subglottal pressure and/or longitudinal vocal-fold tension for the stressed syllable. The speech context used here was necessarily asymmetrical; the /h/ followed an unstressed syllable and initiated a stressed one, to ensure that the /h/ would not be lenited or deleted (Pierrehumbert and Talkin, 1992). Increased P_{sub} as a function of syllable stress should lead to a shorter VOTh (earlier phonation onset), yielding a smaller hysteresis effect. On the other hand, greater longitudinal tension to achieve elevated $f0$ in a stressed syllable should have the effect of postponing voicing. As noted in the Introduction, past studies have found inconsistent results on whether speakers use longitudinal tension as one means of suppressing voicing. For four of the six speakers presented here (F1, F2, F3, F4), the PCA results showed $f0$ loading on the factor(s) that included VOffTh and/or VOTh, suggesting that longitudinal tension of the vocal folds correlated with voice timing in these speakers. The relationship between $f0$ and offset–onset differences is not entirely straightforward, however, a *post hoc* review of each speaker's $f0$ data for stressed and unstressed syllables indicated that neither the absolute $f0$ difference nor the relative $f0$ increase (as a percentage of the unstressed vowel $f0$) showed a clear relationship with the degree of offset–onset differences. A final explanation for the variability in offset–onset differences across speakers might be found in the work of Ní Chasaide and Gobl (1993). In a cross-language study of voice source characteristics around intervocalic consonants, these authors argued that voice offsets are by nature gradual, but that speakers have some leeway in whether they initiate phonation rapidly or more slowly, and further that there may be variation across languages in the manner of voicing onset (rapid vs slow). Interestingly, the five (British) English speakers in their study had greater interspeaker variability in their voice source onset patterns than the speakers of Swedish, German, Italian, and French. Hanson (1997) has also noted considerable cross-speaker variability in voice source characteristics among 22 female speakers, and suggested that women may demonstrate greater interspeaker differences than men with respect to glottal settings. It is a question for further research to what extent the variations in onset–offset differences here reflect language effects, context effects, speaker-specific variations, and/or gender differences. Combining these data with modeling work (cf. Lucero and Koenig, 2005) may help clarify this issue.

Three other general differences between voicing offsets and onsets were observed. First, there was a greater spread of values (higher within-subject SDs) for measures taken at voicing onset than voicing offset. Second, higher r- and F values for voicing onsets in the multiple regressions indicated that the independent measures were more successful in predicting voicing onsets than offsets. Finally, VOffTh and VOTh did not always group with the same variables in the principal components analysis. For example, in speaker F3, VOffTh and vowel quality (/ɪ/-ness) loaded together, whereas the factor that included VOTh had contributions from $f0$ but not from vowel quality. This suggests that, for this speaker, measures of voicing offsets were highly correlated with vowel quality, whereas variation in voicing onsets was more closely related to changes in $f0$. In other words, voicing offsets and onsets may be subject to somewhat different patterns of control within an individual speaker.

B. Cross-speaker similarities

Along with the offset–onset differences described above, some patterns were consistent across speakers. Most notably, hPk and DCOff/On loaded together on at least one factor in the PCA solutions for all speakers. These variables also significantly predicted VOffTh and VOTh in the multiple regression analysis. The consistency of the MR results across speakers implies that, as one might have suspected, all speakers used abduction degree as one way of achieving devoicing.

C. Cross-speaker differences

Apart from the DC flow variables (DCOff, DCOn, hPk), the variables that significantly predicted voice timing in the multiple regressions differed across speakers. Further, the factors that were grouped with VOffTh and VOTh in the principal components analyses differed across speakers. They frequently included $f0$ and the vowel factors, but the direction of vowel effects differed across speakers. This kind of variability is precisely what we would expect in a situation where phonatory timing is determined by a balance among several factors, and where speakers have a variety of options for achieving or sustaining voicing.

The loudness condition also yielded differing effects on the degree of devoicing across speakers. These results are consistent with those of Holmberg and colleagues (1988), who argued that speakers produce changes in loudness via a combination of respiratory and laryngeal adjustments. Vowel and loudness variations also interacted significantly for some speakers in their effects on the frequency of devoicing, suggesting that individual speakers make unique laryngeal and/or respiratory adjustments (or, possibly, show unique patterns of source–tract interaction) in response to changes in supraglottal postures. These contextual changes may affect not only aspects of phonation such as fundamental frequency or voice quality, but the likelihood of phonation itself.

All in all, the results reported here indicate that, while abduction degree (as measured by DCOff/On and hPk) is a major contributing factor to phonation offset and onset, the contributions of other factors vary across speakers. Thus, to some degree, individual speakers appear to develop unique means of achieving voicing and devoicing in running speech. Given the relatively small number of subjects in the current study, we cannot make conclusions about which of the observed patterns may be more common, but the data do give some indication of the possible range of variation among a group of normal female speakers. This variation has implications for understanding the nature of speech production as well as for modeling work that aims to reproduce natural speech behavior and formalize the underlying physical principles of voice production. Since individual speakers satisfy the requirements for voicing in different ways, verifying that a laryngeal model can reproduce the range of normal human vocal behavior requires comparison data from numerous speakers.

As indicated earlier, the current analyses are intended, in part, to provide input to modeling work in which we explore both individual and gender differences in voicing behavior in connected speech. Recently (Lucero and Koenig, 2005), we attempted to reproduce a range of data from male and female speakers using a two-mass model of the vocal folds coupled to a two-tube representation of /ɑ/. The simulation results generally yielded good fits to the data for this vowel context, suggesting that the model does capture a variety of phonatory patterns. The data presented here will permit modeling of phonatory behavior under different supraglottal (vowel) conditions. In future work, we will expand our measured database to include male and female speakers who show little or no devoicing in typical productions of /h/. An additional feature we plan to include in future analyses is voice source measures. These may provide additional information on individual patterns of voicing control, especially in those cases where the current factor analysis did not combine any other variables with the voice timing measures.

ACKNOWLEDGMENTS

This work was supported by NIH Grants DC-00865 and HD-01994 to Haskins Laboratories and by CNPq, Brazil. We are grateful to Anders Löfqvist, Anne Bradlow, and three anonymous reviewers for comments on earlier versions of this manuscript, and we thank the speakers for their participation.

Appendix A

Effect of loudness condition on subglottal pressure (P_{sub}) measures: Means and standard deviations (cm $H_2 0$); ANOVA results; and p-values of post-hoc Scheffé tests. L=Loud; N=Normal; S=Soft.

Speaker	Means (SD) (cm $H_2 0$) L	N	S	ANOVA results F (df=2)	p	Post-hoc p-values L vs N	L vs S	N vs S
F1	6.2 (0.5)	5.4 (1.1)	4.4 (0.3)	81.67	<0.001	<0.001	<0.001	<0.001
F2	7.0 (1.5)	6.9 (1.3)	6.2 (1.7)	6.77	0.001	0.876	0.003	0.015
F3	7.9 (0.8)	6.2 (0.5)	5.9 (0.5)	199.31	<0.001	<0.001	<0.001	0.011
F4	10.6 (0.8)	7.5 (1.3)	4.9 (0.3)	745.46	<0.001	<0.001	<0.001	<0.001
F5	8.3 (2.0)	6.5 (2.3)	7.7 (1.7)	23.95	<0.001	<0.001	0.054	<0.001
F6	9.7 (1.6)	9.0 (2.0)	7.1 (1.8)	43.88	<0.001	0.088	<0.001	<0.001

Appendix B

Effect of loudness condition on hPk values: Means and standard deviations (l/m); ANOVA results; and *p*-values of post-hoc Scheffé tests. *L*=Loud; *N* =Normal; *S*=Soft. Post-hoc results of n/a indicate cases where the post-hoc analyses were not run because the F-test was not significant.

Speaker	Means (SD) (l/m) *L*	*N*	*S*	ANOVA results F (df=2)	*p*	Post-hoc *p*-values *L* vs *N*	*L* vs *S*	*N* vs *S*
F1	57.9 (14.0)	57.6 (16.4)	61.6 (12.3)	1.97	0.140	n/a	n/a	n/a
F2	64.2 (18.4)	41.9 (15.3)	40.5 (14.5)	50.80	<0.001	<0.001	<0.0001	0.863
F3	14.2 (5.7)	15.7 (6.9)	13.7 (4.7)	2.33	0.100	n/a	n/a	n/a
F4	27.8 (9.5)	27.9 (12.5)	55.3 (8.5)	174.20	<0.001	0.998	<0.001	<0.001
F5	23.0 (7.3)	18.0 (7.5)	17.9 (7.6)	11.03	<0.001	<0.001	<0.001	0.998
F6	52.3 (12.7)	43.8 (16.1)	30.4 (7.0)	59.40	<0.001	<0.001	<0.001	<0.001

Appendix C

Effect of vowel on subglottal pressure (P_{sub}) measures: Means and standard deviations (cm H_20); ANOVA results; and *p*-values of post-hoc Scheffé tests.

Speaker	Means (SD) (cm H_20) /ɑ/	/ɪ/	/u/	ANOVA results F (df=2)	*p*	Post-hoc *p*-values /ɑ/ vs /ɪ/	/ɑ/ vs /u/	/ɪ/ vs /u/
FD1	5.1 (0.9)	5.3 (1.0)	5.4 (1.1)	3.10	0.047	0.283	0.034	0.594
FD2	7.0 (1.2)	6.2 (2.1)	6.8 (1.1)	5.79	0.004	0.005	0.659	0.057
FD3	6.4 (0.9)	6.5 (0.9)	6.9 (1.2)	6.81	0.001	0.721	<0.001	0.002
FD4	7.5 (2.5)	7.6 (2.6)	7.7 (2.4)	0.48	0.621	0.903	0.374	0.622
FD5	6.7 (2.0)	7.4 (2.3)	8.4 (1.4)	20.51	<0.001	0.045	<0.001	0.001
FD6	8.4 (2.2)	8.9 (2.2)	8.4 (1.9)	1.82	0.164	0.315	1.000	0.329

Abramson, A. S., and Lisker, L. (**1970**). "Discriminability along the voicing continuum: Cross-language tests," in *Proceedings of the 6th International Congress of Phonetic Sciences, Prague 1967*, edited by B. Hála, M. Romportl, and P. Janota (Academia, Czechoslovak Academy of Sciences, Prague), pp. 569–573.

Abramson, A. S., and Lisker, L. (**1985**). "Relative power of cues: *F*0 shift versus voice timing," in *Linguistic Phonetics: Essays in Honor of Peter Ladefoged*, edited by V. Fromkin (Academic, New York), pp. 25–33.

Alipour-Haghighi, F., and Titze, I. R. (**1991**). "Elastic models of vocal fold tissues," J. Acoust. Soc. Am. **90**, 1326–1331.

Alipour, F., Scherer, R. C., and Finnegan, E. (**1997**). "Pressure-flow relationships during phonation as a function of adduction," J. Voice **11**, 187–194.

Baer, T. (**1975**). "Investigation of phonation using excised larynges," unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Baer, T., Löfqvist, A., and McGarr, N. S. (**1983**). "Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques," J. Acoust. Soc. Am. **73**, 1304–1308.

Behrman, A., and Baken, R. J. (**1997**). "Correlation dimension of electroglottographic data from healthy and pathologic subjects," J. Acoust. Soc. Am. **102**, 2371–2379.

Berke, G. S., Moore, D. M., Monkewitz, P. A., Hanson, D. G., and Gerratt, B. R. (**1989**). "A preliminary study of particle velocity during phonation in an *in vivo* canine model," J. Voice **3**, 306–313.

Berry, D., Herzel, H., Titze, I. R., and Story, B. (**1995**). "Bifurcations in excised larynx experiments," NCVS Status and Progress Report 8, 15–24.

Bickley, C. A., and Stevens, K. N. (**1986**). "Effects of a vocal-tract constriction on the glottal source: Experimental and modelling studies," J. Phonetics **14**, 373–382.

Borden, G. J., and Gay, T. (**1979**). "Temporal aspects of articulatory movements for /s/-stop clusters," Phonetica **36**, 21–31.

Brown, W. S., Jr., and McGlone, R. E. (**1974**). "An aerodynamic and acoustic study of stress in sentence productions," J. Acoust. Soc. Am. **56**, 971–980.

Cattell, R. B. (**1966**). "The scree test for the number of factors," Mult. Behav. Res. **1**, 245–276.

Chan, R. W., Titze, I. R., and Titze, M. R. (**1997**). "Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **101**, 3722–3727.

Childers, D. G., and Lee, C. K. (**1991**). "Vocal quality factors: Analysis, synthesis, and perception," J. Acoust. Soc. Am. **90**, 2394–2410.

Cox, K. A., Alipour, F., and Titze, I. R. (**1999**). "Geometric structure of the human and canine cricothyroid and thyroarytenoid muscles for biomechanical applications," Ann. Otol. Rhinol. Laryngol. **108**, 1151–1158.

Dixit, P., and Brown, W. S., Jr. (**1978**). "Peak magnitudes of supraglottal air pressure associated with affricated and nonaffricated stop consonant productions in Hindi," J. Phonetics **6**, 353–365.

Dixit, R. P., and MacNeilage, P. F. (**1980**). "Cricothyroid activity and control of voicing in Hindi stops and affricates," Phonetica **37**, 397–406.

Dixit, R. P., and Shipp, T. (**1985**). "Study of subglottal air pressure during Hindi stop consonants," Phonetica **42**, 53–77.

Dromey, C., and Ramig, L. O. (**1998**). "Intentional changes in sound-pressure level and rate: Their impact on measures of respiration, phonation, and articulation," J. Speech Lang. Hear. Res. **41**, 1003–1018.

Dromey, C., Stathopoulos, E. T., and Sapienza, C. M. (**1992**). "Glottal airflow and electroglottographic measures of vocal function at multiple intensities," J. Voice **6**, 44–54.

Dunteman, G. H. (**1989**). *Principal Components Analysis* (Sage, Beverly Hills).

Finkelhor, B. K., Titze, I. R., and Durham, P. L. (**1987**). "The effect of viscosity changes in the vocal folds on the range of oscillation," J. Voice **1**, 320–325.

Halle, M., and Stevens, K. N. (**1971**). "A note on laryngeal features," MIT Quart. Prog. Rep. **101**, 198–213.

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

Hertegård, S., and Gauffin, J. (**1995**). "Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography," J. Speech Hear. Res. **38**, 85–100.

Hertegård, S., Gauffin, J., and Lindestad, P. Å. (**1995**). "A comparison of subglottal and intraglottal pressure measurements during phonation," J. Voice **9**, 149–155.

Higgins, M. B., and Saxman, J. H. (**1993**). "Inverse filtered air flow and EGG measures for sustained vowels and syllables," J. Voice **7**, 47–53.

Hirano, M., Kurita, S., and Nakashima, T. (**1983**). "Growth, development, and aging of human vocal folds," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College Hill, San Diego), pp. 22–43.

Hirano, M., Yoshida, T., and Tanaka, S. (**1991**). "Vibratory behavior of human vocal folds viewed from below," in *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, edited by J. Gauffin and B. Hammarberg (Singular, San Diego), pp. 1–6.

Hirose, H., and Niimi, S. (**1987**). "The relationship between glottal opening and the transglottal pressure differences during consonant production," in *Laryngeal Function in Phonation and Respiration*, edited by T. Baer, C. Sasaki, and K. S. Harris (College Hill, Boston), pp. 381–390.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**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.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1989**). "Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch," J. Voice **3**, 294–305.

Holmberg, E. B., Hillman, R. E., Perkell, J. J., and Gress, C. (**1994**). "Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings," J.

Speech Hear. Res. **37**, 484–495.
Ishizaka, K., and Flanagan, J. L. (**1972**). "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233–1268.
Jackson, J. E. (**1991**). *A User's Guide to Principal Components* (Wiley, New York).
Jiang, J. J., and Titze, I. R. (**1993**). "A methodological study of hemilaryngeal phonation," Laryngoscope **103**, 872–882.
Jiang, J., Raviv, J. R., and Hanson, D. G. (**2001**). "Comparison of the phonation-related structures among pig, dog, white-tailed deer, and human larynges," Ann. Otol. Rhinol. Laryngol. **110**, 1120–1125.
Johnson, K., Ladefoged, P., and Lindau, M. (**1993**). "Individual differences in vowel production," J. Acoust. Soc. Am. **94**, 701–714.
Kagaya, R., and Hirose, H. (**1975**). "Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants," Ann. Bulletin, Res. Inst. Logop. and Phoniat. Univ. Tokyo **9**, 27–46.
Kahane, J. C. (**1987**). "Connective tissue changes in the larynx and their effects on voice," J. Voice **1**, 27–30.
Kahane, J. C. (**1988**). "Age related changes in the human cricoarytenoid joint," in *Vocal Physiology: Voice Production, Mechanisms and Functions*, edited by O. Fujimura (Raven, New York), Vol. **2**, pp. 145–157.
Kim, M. J., Hunter, E. J., and Titze, I. R. (**2004**). "Comparison of human, canine, and ovine laryngeal dimensions," Ann. Otol. Rhinol. Laryngol. **113**, 60–68.
Kitzing, P., and Sonesson, B. (**1974**). "A photoglottographic study of the female vocal folds during phonation," Folia Phoniatr. **26**, 138–149.
Klatt, D. H., Stevens, K. N., and Mead, J. (**1968**). "Studies of articulatory activity and airflow during speech," in *Sound Production in Man*, edited by A. Bouhuys (New York Academy of Sciences, New York), pp. 42–55.
Koenig, L. L. (**2000**). "Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds," J. Speech Lang. Hear. Res. **43**, 1211–1228.
Lindqvist, J. (**1972**). "Laryngeal articulation studied on Swedish subjects," Quart. Status & Prog. Rep. (Spch. Transm. Lab., Royal Inst. of Tech., Stockholm) **2–3**, 10–27.
Lisker, L., and Abramson, A. S. (**1964**). "A cross-language study of voicing in initial stops: Acoustical measurements," Word **20**, 384–422.
Lisker, L., and Abramson, A. S. (**1970**). "The voicing dimension: Some experiments in comparative phonetics," in *Proceedings of the 6th International Congress of Phonetic Sciences, Prague 1967*, edited by B. Hála, M. Romportl, and P. Janota (Academia, Czechoslovak Academy of Sciences, Prague), pp. 563–567.
Lisker, L., Sawashima, M., Abramson, A. S., and Cooper, F. S. (1970). "Cinegraphic observations of the larynx during voiced and voiceless stops," Haskins Labs. Status Repts. on Speech Res. **21/22**, 201–210.
Löfqvist, A. (**1975**). "A study of subglottal pressure during the production of Swedish stops," J. Phonetics **3**, 175–189.
Löfqvist, A. (**1992**). "Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants," Lang Speech **35**, 15–28.
Löfqvist, A., Baer, T., McGarr, N. S., and Story, R. S. (**1989**). "The cricothyroid muscle in voicing control," J. Acoust. Soc. Am. **85**, 1314–1321.
Löfqvist, A., Carlborg, B., and Kitzing, P. (**1982**). "Initial validation of an indirect measure of subglottal pressure during vowels," J. Acoust. Soc. Am. **72**, 633–635.
Löfqvist, A., Koenig, L. L., and McGowan, R. S. (**1995**). "Vocal tract aerodynamics in /aCa/ utterances: Measurements," Speech Commun. **16**, 49–66.
Löfqvist, A., and McGowan, R. S. (**1992**). "Influence of consonantal environment on voice source aerodynamics," J. Phonetics **20**, 93–110.
Lucero, J. C. (**1995**). "The minimum lung pressure to sustain vocal fold oscillation," J. Acoust. Soc. Am. **98**, 779–784.
Lucero, J. C. (**1999**). "A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset–offset," J. Acoust. Soc. Am. **105**, 423–431.
Lucero, J. C., and Koenig, L. L. (**2005**). "Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control," J. Acoust. Soc. Am. **117**, 1362–1372.
Matsushita, H. (**1975**). "The vibratory mode of the vocal folds in the excised larynx," Folia Phoniatr. **26**, 7–18.
McGlone, R. E., and Shipp, T. (**1972**). "Comparison of subglottal air pressures associated with /p/ and /b/," J. Acoust. Soc. Am. **51**, 664–665.
Munhall, K. G., Löfqvist, A., and Kelso, J. A. S. (**1994**). "Lip–larynx coordination in speech: Effects of mechanical perturbations to the lower lip," J. Acoust. Soc. Am. **95**, 3605–3616.
Ní Chasaide, A., and Gobl, C. (**1993**). "Contextual variation of the vowel voice source as a function of adjacent consonants," Lang Speech **36**, 303–330.
Ohala, J. J. (**1990**). "Respiratory activity in speech," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht), pp. 23–54.
Pierrehumbert, J., and Talkin, D. (**1992**). "Lenition of /h/ and glottal stop," in *Papers in Laboratory Phonology II: Segment, Gesture, Tone*, edited by G. Docherty and D. R. Ladd (Cambridge University Press, Cambridge), pp. 90–117.
Plant, R. L., Freed, G. L., and Plant, R. E. (**2004**). "Direct measurement of onset and offset phonation threshold pressure in normal subjects," J. Acoust. Soc. Am. **116**, 3640–3646.
Price, P. J. (**1989**). "Male and female voice source characteristics: Inverse filtering results," Speech Commun. **8**, 261–277.
Raphael, L. J., and Bell-Berti, F. (**1975**). "Tongue musculature and the feature of tension in English vowels," Phonetica **32**, 61–73.
Rothenberg, M. (**1973**). "A new inverse-filtering technique for deriving the glottal air flow waveform during voicing," J. Acoust. Soc. Am. **53**, 1632–1645.
Saito, S., Fukuda, H., Kitahara, S., Isogai, Y., Tsuzuki, T., Muta, H., Takayama, E., Fujioka, T., Kokawa, N., and Makino, K. (**1983**). "Pellet tracking in the vocal fold while phonating—experimental study using canine larynges with muscle activity," in *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*, edited by I. R. Titze and R. C. Scherer (Denver Center for the Performing Arts, Denver), pp. 169–182.
Smitheran, J. R., and Hixon, T. J. (**1981**). "A clinical method for estimating laryngeal airway resistance during vowel production," J. Speech Hear Disord. **46**, 138–146.
Stevens, K. N. (**1977**). "Physics of laryngeal behavior and larynx modes," Phonetica **34**, 264–279.
Timcke, R., von Leden, H., and Moore, P. (**1958**). "Laryngeal vibrations: Measurement of the glottic wave. I. The normal vibratory cycle," AMA Arch. Otolaryngol. **68**, 1–19.
Titze, I. R. (**1988**). "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552.
Titze, I. R. (**1989**). "Physiologic and acoustic differences between male and female voices," J. Acoust. Soc. Am. **85**, 1699–1707.
Titze, I. R. (**1992**). "Phonation threshold pressure: A missing link in glottal aerodynamics," J. Acoust. Soc. Am. **91**, 2926–2935.
Titze, I. R., Jiang, J. J., and Hsiao, T.-Y. (**1993**). "Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration," Ann. Otol. Rhinol. Laryngol. **102**, 58–62.
Titze, I. R., Schmidt, S. S., and Titze, M. R. (**1995**). "Phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **97**, 3080–3084.
van den Berg, J. (**1958**). "Myoelastic-aerodynamic theory of voice production," J. Speech Hear. Res. **1**, 227–244.
van den Berg, J., and Tan, T. S. (**1959**). "Results of experiments with human larynxes," Pract. Otorhinolaryngol. (Basel) **21**, 425–450.
Whalen, D. H., and Levitt, A. G. (**1995**). "The universality of intrinsic *f*0 of vowels," J. Phonetics **23**, 349–366.
Yumoto, E., Kadota, Y., and Kurokawa, H. (**1993**). "Infraglottic aspect of canine vocal fold vibration: Effect of increase of mean airflow rate and lengthening of vocal fold," J. Voice **7**, 311–318.