

# Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy<sup>a)</sup>

Daryush D. Mehta<sup>b)</sup>

Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, Massachusetts 02114

Matías Zañartu<sup>c)</sup>

Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

Thomas F. Quatieri

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02421

Dimitar D. Deliyski<sup>d)</sup>

Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229

Robert E. Hillman<sup>e)</sup>

Center for Laryngeal Surgery and Voice Rehabilitation & MGH Institute of Health Professions, Massachusetts General Hospital, Boston, Massachusetts 02114

(Received 16 December 2010; revised 29 September 2011; accepted 6 October 2011)

Vocal fold vibratory asymmetry is often associated with inefficient sound production through its impact on source spectral tilt. This association is investigated in both a computational voice production model and a group of 47 human subjects. The model provides indirect control over the degree of left–right phase asymmetry within a nonlinear source–filter framework, and high-speed videoendoscopy provides *in vivo* measures of vocal fold vibratory asymmetry. Source spectral tilt measures are estimated from the inverse-filtered spectrum of the simulated and recorded radiated acoustic pressure. As expected, model simulations indicate that increasing left–right phase asymmetry induces steeper spectral tilt. Subject data, however, reveal that none of the vibratory asymmetry measures correlates with spectral tilt measures. Probing further into physiological correlates of spectral tilt that might be affected by asymmetry, the glottal area waveform is parameterized to obtain measures of the open phase (open/plateau quotient) and closing phase (speed/closing quotient). Subjects' left–right phase asymmetry exhibits low, but statistically significant, correlations with speed quotient ( $r = 0.45$ ) and closing quotient ( $r = -0.39$ ). Results call for future studies into the effect of asymmetric vocal fold vibration on glottal airflow and the associated impact on voice source spectral properties and vocal efficiency. © 2011 Acoustical Society of America. [DOI: 10.1121/1.3658441]

PACS number(s): 43.70.Gr, 43.70.Bk, 43.70.Dn, 43.70.Jt [DAB]

Pages: 3999–4009

## I. INTRODUCTION

During sustained vowel phonation, the vocal folds are set into vibration by a combination of muscle tension and aerodynamic forces (van den Berg, 1958). Vibratory phase

asymmetries between the left and right vocal folds have been observed in speakers with normal voices and with voice disorders (Haben *et al.*, 2003; Bonilha *et al.*, 2008, 2011), with various factors purported to influence asymmetry within speakers, including subglottal pressure (Berry *et al.*, 1996; Maunsell *et al.*, 2006; Murugappan *et al.*, 2009), fundamental frequency (Maunsell *et al.*, 2006), vocal fold mass and stiffness (Steinecke and Herzel, 1995), and vocal loading (Lohscheller *et al.*, 2008a).

The goals of the current work are motivated by the clinical need for systematic studies to describe and develop acoustic correlates of vocal fold vibratory asymmetry to potentially aid clinicians in the effective management of voice disorders (Mehta and Hillman, 2008). Voice specialists make critical diagnostic, medical, therapeutic, and surgical decisions based on coupling visual observations of vocal fold tissue motion with auditory-perceptual assessments of voice quality (Zeitels *et al.*, 2007). Although clinical experience indicates that this

<sup>a)</sup>Portions of this work were presented at the 9th International Conference on Advances in Quantitative Laryngology in Erlangen, Germany, in September 2010, and in the doctoral dissertation of the first author (Mehta, 2010).

<sup>b)</sup>Author to whom correspondence should be addressed. Also at: School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138. Electronic mail: daryush.mehta@alum.mit.edu

<sup>c)</sup>Portions of this work were performed at School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907.

<sup>d)</sup>Also at Departments of Otolaryngology—Head and Neck Surgery, Pediatrics, & Communication Sciences and Disorders, University of Cincinnati, Cincinnati, Ohio 45229.

<sup>e)</sup>Also at Surgery and Health Sciences & Technology, Harvard Medical School, Boston, Massachusetts 02115.

approach is generally applicable, it is inherently limited to case-by-case observations, and visual judgments of vocal fold vibratory patterns might not adequately reflect changes in objective measures of the acoustic signal (Haben *et al.*, 2003).

A handful of studies have found presumably strong links between the presence of asymmetric vocal fold vibration and degradations in voice quality. Using acoustic recordings and high-speed videokymography from four voice patients, the presence of left–right phase asymmetry has been associated with the auditory perception of roughness (Verdonck-de Leeuw *et al.*, 2001). In a separate study of 22 subjects with voice disorders, ratings of roughness and breathiness were shown to be statistically different for subjects exhibiting symmetric versus asymmetric vocal fold vibration (Niimi and Miyaji, 2000). These results suggest that acoustic correlates of vocal fold vibratory phase asymmetry should include acoustic perturbation measures (related to the perception of roughness) and the harmonics-to-noise ratio (related to the perception of breathiness).

In a high-speed videoendoscopy (HSV) study of 14 voice patients, the average degree of vocal fold vibratory asymmetry did not correlate statistically with levels of acoustic jitter, shimmer, or harmonics-to-noise ratio during sustained vowel phonation (Mehta *et al.*, 2010). Instead, a significant amount of the variation in acoustic jitter was accounted for by the standard deviation of the degree of asymmetry. Further, the harmonics-to-noise ratio was within normal ranges in these subjects, largely due to the maintenance of glottal closure and vibratory periodicity. Thus, increasing degrees of asymmetric vocal fold vibration were not necessarily indicators of elevated levels of acoustic perturbation or noise energy.

In a classification of high-speed videokymography recordings from 45 voice patients, different types of vibratory asymmetry were described qualitatively from the lateral displacement patterns exhibited by the left and right vocal folds (Švec *et al.*, 2007). Categories of vocal fold vibratory asymmetry included left–right amplitude asymmetry, left–right phase asymmetry, left–right frequency differences, and axis shift during glottal closure. Objective methods for quantifying the various types of asymmetry have been suggested in previous work (Deliyski and Petrushev, 2003; Qiu *et al.*, 2003; Bonilha *et al.*, 2008; Mehta *et al.*, 2010).

HSV methods have also enabled the visualization and quantification of vocal fold vibratory asymmetries and the glottal area waveform in *ex vivo* configurations. Evidence from excised canine larynx experiments suggests that the presence of vocal fold vibratory asymmetry accompanies an increase in closing quotient, a decrease in closing velocity, and a peak reduction and flattening of the glottal area waveform (Khosla *et al.*, 2008; Khosla *et al.*, 2009). The modified source excitation yielded a steeper spectral tilt of the acoustic pressure waveform at the laryngeal exit. Also, the presence of vibratory phase asymmetry, induced via unilateral vocal fold scarring, has been shown to accompany a reduction in flow separation vortices (Murugappan *et al.*, 2009).

In the current study, the physiological observations are explored in a computational model of voice production and in a group of subjects with and without voice disorders. In a widely applied model of the vocal folds, a set of differential

equations simulates self-sustained oscillations by coupling Bernoulli airflow mechanisms to a lumped-element parameterization of the vocal folds (Ishizaka and Flanagan, 1972). Asymmetric vocal fold vibratory regimes have been explored in adaptations of this two-mass model, in which each vocal fold is represented by separate superior and inferior mass elements to allow for a vertical phase difference (mucosal wave), and left–right vibratory phase asymmetry is allowed through modifications of the mass, stiffness, and damping properties of each vocal fold (Ishizaka and Isshiki, 1976; Steinecke and Herzel, 1995).

The model proposed by Steinecke and Herzel (1995) has been repeatedly used to study asymmetric behavior during phonation (Schwarz *et al.*, 2006; Wurzbacher *et al.*, 2006; Zhang *et al.*, 2006) and whose behavior has been validated by more complex models that use Navier–Stokes flow solvers (Tao *et al.*, 2007; Xue *et al.*, 2010). Although alternative representations, such as the body-cover model (Story and Titze, 1995), may offer an enhanced feature set, they do not provide simple control over left–right vibratory asymmetry and have not been consistently applied to study asymmetric vocal fold vibration. In addition, parametric glottal airflow characterizations, such as the LF model (Fant, 1997), are applicable, but do not have available the physiologically based property of self-sustained vocal fold vibration with asymmetries and may yield nonphysiological oscillatory conditions.

Thus, the core Steinecke and Herzel (1995) model is chosen for the current study to allow for comparison with prior studies. Additional model features, however, are implemented that are fundamental to the present acoustic investigation. In particular, nonlinear source–filter coupling between subglottal and supraglottal tracts is incorporated to account for acoustic coupling that has been shown to significantly affect source properties and system dynamics (Story and Titze, 1995; Zañartu *et al.*, 2007; Titze *et al.*, 2008; Titze, 2008). This coupling allows for observing the well-documented pulse skewing of the glottal airflow waveform, which is critical for determining spectral tilt and other acoustic characteristics (Doval and d’Alessandro, 2006). Acoustic coupling is also incorporated in the vocal fold dynamics to keep the physical assumptions consistent.

The purpose of the current work is to better understand the relationships between asymmetric vocal fold vibration and spectral measures of the radiated acoustic pressure waveform. Simulations of asymmetric vocal fold vibration are performed using a mathematical model of the vocal folds that offers indirect control over the vibratory phase asymmetry between the left and right vocal folds. Data from HSV recordings of human subjects are then compared with results from the model simulations. Correlational analysis, in addition to illustrative case studies, is performed to document the impact of phase asymmetry on properties of the glottal area waveform and spectral measures of the radiated acoustic pressure waveform.

## II. METHODS

### A. Computational model of voice production

Figure 1 schematizes the proposed computational model of Steinecke and Herzel (1995), which is extended to allow

for nonlinear acoustic coupling. Each vocal fold (subscript  $\alpha = l$  for left and  $\alpha = r$  for right) is represented by lower and upper coupled oscillators with masses  $m_{1\alpha}$  and  $m_{2\alpha}$ , respectively. Mechanical spring constants  $k_{1\alpha}$ ,  $k_{2\alpha}$ , and  $k_{c\alpha}$ , and damping parameters  $r_{1\alpha}$  and  $r_{2\alpha}$ , describe viscoelastic tissue properties. Additional spring constants  $c_{1\alpha}$  and  $c_{2\alpha}$  (not shown in Fig. 1) are activated during collision of the lower and upper elements, respectively. The vocal fold displacements from rest  $x_{1\alpha}$  and  $x_{2\alpha}$  are restricted to the horizontal (mediolateral) axis, and  $x_{01\alpha}$  and  $x_{02\alpha}$  are the rest displacements of the lower and upper elements, respectively.

Left vocal fold parameters are set to standard values that are known to behave appropriately to model modal phonation with a subglottal pressure of  $\sim 8$  cm H<sub>2</sub>O (Steinecke and Herzog, 1995, p. 1876). As in the original model, asymmetric vocal fold vibration is produced by applying a  $Q$  factor to scale certain parameter values of the opposing vocal fold:

$$\begin{aligned} m_{1r} &= m_{1l}/Q, & m_{2r} &= m_{2l}/Q, \\ k_{1r} &= Q \cdot k_{1l}, & k_{2r} &= Q \cdot k_{2l}, \\ c_{1r} &= Q \cdot c_{1l}, & c_{2r} &= Q \cdot c_{2l}, \\ k_{cr} &= Q \cdot k_{cl}, \\ r_{1r} &= r_{1l}, & r_{2r} &= r_{2l}, \\ x_{01r} &= x_{01l}, & x_{02r} &= x_{02l}. \end{aligned} \quad (1)$$

This configuration was originally used to model a superior laryngeal nerve paralysis and is applied in the current study with the intent of representing imbalances in vocal fold tissue properties due to pathologies such as polyps and glottic cancer. Thus, changes in  $Q$  simulate different degrees of left–right phase asymmetry.

Four equations of motion describe the oscillation of the mass elements (time dependencies omitted for clarity):

$$\begin{aligned} m_{1\alpha} \ddot{x}_{1\alpha} + r_{1\alpha} \dot{x}_{1\alpha} + k_{1\alpha} x_{1\alpha} + \Theta(-a_1) c_{1\alpha} [a_1/2l] \\ + k_{c\alpha} (x_{1\alpha} - x_{2\alpha}) = l d_1 P_1, \end{aligned} \quad (2a)$$

$$\begin{aligned} m_{2\alpha} \ddot{x}_{2\alpha} + r_{2\alpha} \dot{x}_{2\alpha} + k_{2\alpha} x_{2\alpha} + \Theta(-a_2) c_{2\alpha} [a_2/2l] \\ + k_{c\alpha} (x_{2\alpha} - x_{1\alpha}) = l d_2 P_2, \end{aligned} \quad (2b)$$

where  $P_1$  and  $P_2$  are the time-varying intraglottal pressures between the lower and upper mass elements, respectively;

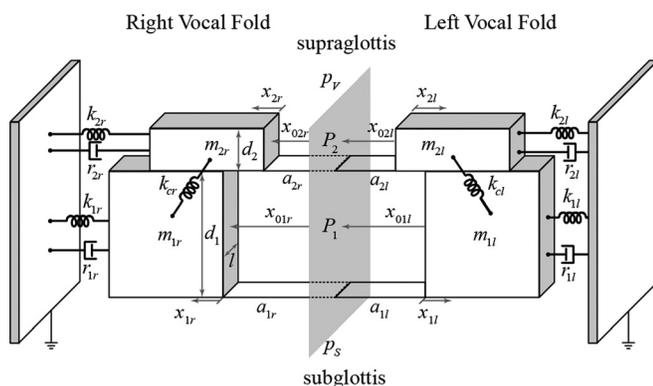


FIG. 1. Schematic diagram of the lumped-element vocal fold model. Vertical gray plane represents the glottal midplane. See the text for definitions of variables.

$a_1 = a_{1l} + a_{1r}$ , and  $a_2 = a_{2l} + a_{2r}$ , are the time-varying area functions between the lower and upper mass elements, respectively;  $d_1$  and  $d_2$  are the inferior–superior dimensions of the lower and upper mass elements, respectively; and  $l$  is the anterior–posterior length of the glottis.  $\Theta(x) = \tanh[50(x/x_0)]$  when  $x > 0$ , and  $\Theta(x) = 0$  otherwise ( $x_0 = 0.05$  cm<sup>2</sup>). Thus, when  $a_1$  and  $a_2$  are negative (i.e., during collision), the additional spring constants  $c_{1\alpha}$  and  $c_{2\alpha}$  play a role in the equations of motion.

To allow for nonlinear acoustic coupling, acoustic pressures are added to the assumed Bernoulli regime below the narrowest part of the glottis and to the jet regime above the glottal narrowing (Story and Titze, 1995). The equations for the intraglottal pressures  $P_1$  and  $P_2$  are

$$P_1 = [P_S + p_S - (P_S + p_S - p_V)(a_{\min}/a_1^2)] \Theta(a_1), \quad (3a)$$

$$P_2 = p_V \Theta(a_2), \quad (3b)$$

where  $P_S$  is the static subglottal pressure,  $p_S$  is the acoustic subglottal pressure,  $p_V$  is the acoustic supraglottal pressure, and  $a_{\min} = \max[0, \min(a_{1l}, a_{2l}) + \min(a_{1r}, a_{2r})]$ . The oscillation amplitudes  $x_{1\alpha}$  and  $x_{2\alpha}$  are solved for using the explicit Runge–Kutta RK5(4) formula (Dormand and Prince, 1980) with a step size corresponding to a 70 kHz sampling rate. The applied initial conditions— $x_{1\alpha}(0) = \dot{x}_{1\alpha}(0) = 0.1$  and  $x_{2\alpha}(0) = \dot{x}_{2\alpha}(0) = 0$ —force nontrivial solutions.

Nonlinear source–filter interactions are incorporated by coupling vocal fold dynamics to wave reflection analog models of the supraglottal and subglottal tracts. The wave reflection analog model is a time-domain description of the propagation of one-dimensional planar acoustic waves through a series of uniform cylindrical tubes. Per Titze (2008) the acoustic pressures  $p_S$  and  $p_V$  are allowed to interact with and affect both glottal airflow (level 1 interaction) and vocal fold tissue motion (level 2 interaction). An interactive airflow model at the glottis is applied (Titze, 2002, Eq. A53), which was originally designed to describe nonlinear coupling using a transmission line model (Titze, 1984, Sec. III C).

Figure 2 illustrates the tract geometries of the voice production model in axial sections  $A = 0.25$  cm in width. The 70 kHz sampling rate is derived from  $C/(2A)$ , where the speed of sound  $C = 350$  m/s. The model includes a radiation impedance (Story and Titze, 1995) and different loss factors for the subglottal and supraglottal tracts (Zañartu et al., 2007). The vocal tract area function is taken from 3D cine-MRI data of an adult male sustaining the /e/ phoneme (Takemoto et al., 2006). The subglottal area function is adapted from respiratory system measurements of human

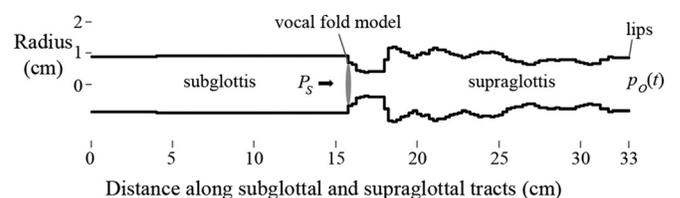


FIG. 2. Subglottal and supraglottal tract geometry in the voice production model. For illustration, the lower half of the plot mirrors the top half.  $P_S$  = static subglottal pressure,  $p_O(t)$  = acoustic pressure.

cadavers (Weibel, 1963) and includes the trachea, bronchi, and a resistive termination impedance (zeroth and first airway generations).

## B. Human subject data

### 1. Subject selection

Data are collected from 47 subjects: 40 individuals with voice disorders (24 male, 16 female) and 7 speakers with no history of voice disorders (4 male, 3 female). Six of the subjects with voice disorders participated before and after voice surgery, yielding a total of 53 recordings across both subject groups. The average age of subjects with voice disorders is 55 years, with a range of 19–85 years. Speakers with normal voices are 33 years old on average, with a range of 20–52 years.

### 2. Data acquisition

Subjects underwent laryngeal HSV using transoral rigid endoscopy and were instructed to sustain the vowel /i/ at a comfortable pitch and loudness for four seconds. HSV recordings are acquired with a color high-speed video camera (Phantom v7.3; Vision Research, Inc., Wayne, NJ). Video frame rates are set at either 4,000 Hz (4 subjects) or 6,250 Hz (43 subjects) with maximum integration time. The differences in sampling rate did not have a significant effect on outcome measures because measures are averaged across phonatory segments. Spatial resolution is  $\sim 0.05$  mm per pixel in the 320-by-352 pixel array.

The acoustic signal is recorded using a head-mounted condenser microphone (MKE104, Sennheiser electronic GmbH, Wennebostel, Germany). The microphone is positioned  $\sim 4$  cm from the lips at a  $45^\circ$  azimuth. The analog signal is digitized at a 100 kHz sampling rate and 16-bit quantization. Time synchronization of the HSV recordings and the acoustic signal to within  $11 \mu\text{s}$  is enabled by a camera synchronization signal and a common clock source to the camera and data acquisition board (MiDAS DA, Xcitex Corporation, Cambridge, MA). The microphone signal is time-shifted into the past by  $600 \mu\text{s}$  relative to the HSV recording to compensate for the larynx-to-microphone acoustic propagation time.

## C. Data analysis

### 1. Preprocessing of model waveforms

The duration of each model simulation is 500 ms, with the initial 250 ms discarded to neglect transient effects. Waveforms are resampled to an 8 kHz sampling rate. Figure 3 illustrates the derivation of lateral displacement waveforms  $x_L$  and  $x_R$  of the left and right vocal folds, respectively, which are defined when the minimum glottal area  $a_{\min}$  is positive:

$$x_L = \begin{cases} \min(x_{1L}, x_{2L}), & a_{\min} > 0 \\ \text{undefined} & \text{otherwise} \end{cases}, \quad (4a)$$

$$x_R = \begin{cases} \min(x_{1R}, x_{2R}), & a_{\min} > 0 \\ \text{undefined} & \text{otherwise} \end{cases}, \quad (4b)$$

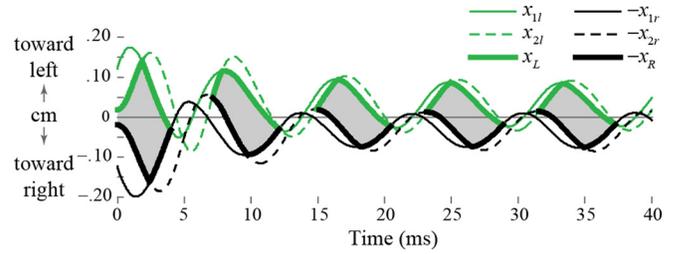


FIG. 3. (Color online) Illustration of how lateral displacement waveforms  $x_L$  and  $x_R$  (of the left and right vocal folds, respectively) are derived from model outputs  $x_{1\alpha}$  and  $x_{2\alpha}$  ( $\alpha = l, r$ ) with  $Q = 0.8$ . Gray shading indicates times of positive glottal area.

where

$$x_{1L} = \frac{a_{01}}{2l} + x_{1l}, \quad x_{2L} = \frac{a_{02}}{2l} + x_{2l}, \quad (5a)$$

$$x_{1R} = \frac{a_{01}}{2l} + x_{1r}, \quad x_{2R} = \frac{a_{02}}{2l} + x_{2r}. \quad (5b)$$

Notably, the onsets of glottal opening and closure do not necessarily occur at the glottal midplane (gray plane in Fig. 1), and  $a_{\min}$  assumes a rectangular glottal aperture. The acoustic voice signal is the radiated acoustic pressure  $p_o(t)$  at the “lips” of the model (see Fig. 2).

### 2. Preprocessing of subject data

Stable phonatory segments between 320 and 400 ms in duration are selected from each HSV recording and corresponding acoustic voice signal. HSV processing follows previously described methodologies (Mehta *et al.*, 2010; Mehta *et al.*, 2011). Briefly, motion artifacts are compensated for so that edge detection is equivalent to motion tracking of the vocal fold edge closest to the midline (Deliyski, 2005). The glottal midline (anterior–posterior axis) is defined interactively on the first image capturing maximum vocal fold displacement. Endpoints of the midline indicate the anterior commissure and the posterior end of the musculo-membranous glottis. All HSV images are cropped and rotated such that the glottal midline is oriented vertically.

Lateral displacement waveforms  $x_L$  and  $x_R$  are derived from a digital kymogram (DKG) taken halfway between the endpoints of the glottal midline. The DKG is converted to an 8 bit monochromatic space by keeping red-channel information. A user-defined intensity threshold segments the relatively dark glottis from tissue regions. Upper and lower edges of the DKG segmentation define  $x_L$  and  $x_R$ , respectively. As in the model, the lateral displacement waveforms are undefined during glottal closure. The glottal area waveform  $a_{\min}$  is derived from the full-frame HSV images through threshold-based glottal segmentation using an intensity threshold for the entire phonatory segment (Mehta *et al.*, 2011).

A phonovibrogram (PVG) is generated to aid in visualizing vocal fold displacements at different positions along the anterior–posterior length of the glottis (Lohscheller *et al.*, 2008b). The PVG provides qualitative information

regarding the uniformity of left–right phase and amplitude asymmetries in the subject data. PVGs are not shown for the model outputs because the rectangular glottal aperture inherently yields no anterior–posterior vibratory differences.

### 3. Measures of vocal fold vibratory asymmetry

Figure 4(A) illustrates the parameterization of the lateral displacement waveforms to obtain measures of left–right phase asymmetry, left–right amplitude asymmetry, and axis shift during closure, as developed in previous work (Mehta *et al.*, 2011).

Left–right phase asymmetry PA quantifies the normalized phase delay between the left and right vocal folds (Bonilha *et al.*, 2008; Lohscheller *et al.*, 2008b):

$$PA = \frac{n_R - n_L}{OP}, \quad (6)$$

where  $n_R$  ( $n_L$ ) is the time index at maximum lateral displacement of the right (left) and OP is the duration of the open phase. The normalization factor OP is different from that in a previous study (Mehta *et al.*, 2011) because of the current focus on acoustic effects generated during the open phase.

Left–right amplitude asymmetry AA quantifies the relative peak-to-peak displacements of the left and right vocal folds (Qiu *et al.*, 2003):

$$AA = \frac{A_L - A_R}{A_L + A_R}, \quad (7)$$

where  $A_L$  ( $A_R$ ) is the peak-to-peak lateral displacement of the left (right) vocal fold.

The axis shift AS is the mediolateral distance traveled by the vocal folds during glottal closure (Švec *et al.*, 2007):

$$AS = \frac{x_o - x_c}{w}, \quad (8)$$

where  $x_o$  is the mediolateral position of the vocal folds at the onset of glottal opening,  $x_c$  is the mediolateral position of the vocal folds at glottal closure of the previous cycle, and  $w$  is the maximum glottal width.

### 4. Glottal area waveform measures

Figure 4(B) illustrates the parameterization of the glottal area waveform that is motivated by studies on acoustic correlates of glottal characteristics using inverse-filtered airflow waveforms (Holmberg *et al.*, 1988). Open quotient OQ (ratio of open phase duration OP to period  $P$ ), speed quotient SQ (ratio of opening phase duration  $a$  to closing phase duration  $b$ ), and closing quotient CIQ (ratio of closing phase duration  $b$  to period  $P$ ) are computed for each cycle:

$$OQ = \frac{OP}{P}, \quad SQ = \frac{a}{b}, \quad CIQ = \frac{b}{P}. \quad (9)$$

An additional measure, plateau quotient PQ, describes the peakiness of the glottal area waveform:

$$PQ = \frac{PP}{OP}, \quad (10)$$

where PP is the plateau phase duration, defined when the glottal area is greater than 95% of its maximum. The inclusion of PQ is motivated by observations of flattened peaks of the glottal area waveform in the asymmetric vibration of excised canine larynges (Khosla *et al.*, 2008; Khosla *et al.*, 2009; Murugappan *et al.*, 2009).

### 5. Spectral measures of acoustic pressure

Acoustic measures commonly linked to voice quality are estimated from the magnitude spectrum of the radiated acoustic pressure waveform. An inverse filtering method is applied to compensate harmonic magnitudes for the effects of the vocal tract transfer function to obtain source-related spectral measures (Iseli *et al.*, 2007). Estimates of the frequency and bandwidth of the first three formants (Boersma and Weenink, 2009) are used to derive compensated harmonic magnitudes (denoted by asterisks), which are labeled by harmonic number (e.g.,  $H1^*$ ) or by nearest formant number (e.g.,  $A1^*$ ).

The following acoustic spectral measures, in decibels, are computed due to their relations to source-related characteristics:  $H1^*-H2^*$ , a correlate of open quotient (Holmberg *et al.*, 1995);  $H1^*-A1^*$ , a correlate of the first formant bandwidth and energy efficiency (Hanson and Chuang, 1999); and  $H1^*-A3^*$ , an estimate of spectral tilt (Hanson and Chuang, 1999). An alternative measure of spectral tilt,  $TL^*$ , is computed as the linear regression slope, in decibels per octave, over the first eight compensated harmonic magnitudes ( $H1^*$ ,  $H2^*$ , ...,  $H8^*$ ).  $TL^*$  mirrors a spectral tilt measure applied in excised larynx work (Murugappan *et al.*, 2009). These spectral measures assume a linear source–filter

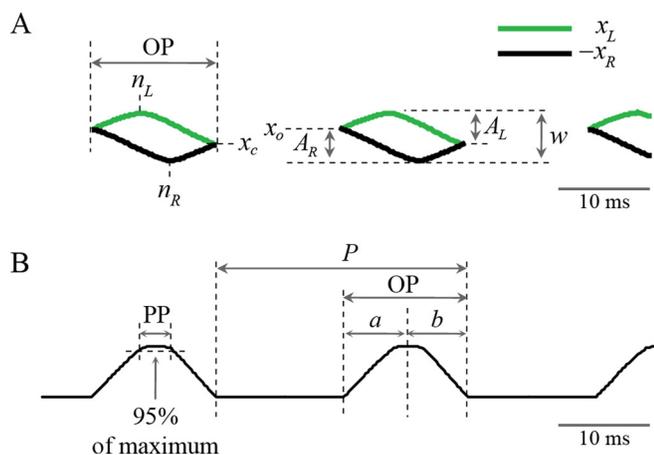


FIG. 4. (Color online) Parameterization of (A) lateral displacement waveforms and (B) glottal area waveforms to obtain vocal fold vibratory measures.  $x_L$ ,  $x_R$  = lateral displacement of left/right vocal fold; OP = open phase;  $n_L$ ,  $n_R$  = time of maximal left/right vocal fold displacement;  $x_c$ ,  $x_o$  = mediolateral position of vocal folds at onset of glottal closure/opening;  $A_L$ ,  $A_R$  = peak-to-peak lateral displacement of left/right vocal fold;  $w$  = maximum glottal width; PP = plateau phase;  $P$  = period; and  $a$ ,  $b$  = opening/closing phase.

framework, in which source harmonic magnitudes typically exhibit a monotonic decay with increasing frequency.

## 6. Statistical analysis

Measures of the lateral displacement and glottal area waveforms are averaged over all the cycles in the phonatory segment. Pearson's correlation coefficient  $r$  describes the pairwise relationships among measures computed from lateral displacement waveforms, glottal area waveforms, and radiated acoustic pressure waveforms. Reported correlations are statistically significant at 95% confidence levels. Scatter plots display marginal and joint distributions of the measures.

## III. RESULTS

### A. Model simulations

#### 1. Control over left–right phase asymmetry

Figure 5 shows the effect of  $Q$  on left–right phase asymmetry PA and left–right amplitude asymmetry AA over the range  $0.77 \leq Q \leq 1$ . Given the linear relationship between  $Q$  and PA, subsequent model analysis is restricted to this regime so that changes in  $Q$  map linearly to changes in PA. The model yields a nonlinear relationship between  $Q$  and PA when  $Q$  is less than 0.77. AA exhibits nonlinear effects with changes in  $Q$  but remains low (less than 7.2%) for all values of  $Q$  in the range of interest. Note that any situation with  $Q > 1$  can be transformed to a case with  $Q < 1$ .

Figure 6 displays waveforms from two model simulations with extreme  $Q$  values. Selected measures computed with  $Q = 1$ : PA = 0%, AA = 0%, CIQ = 37.3%, and TL\* = -4.6 dB/octave. In contrast, with  $Q = 0.77$ , the right vocal fold masses increase and stiffness parameters decrease relative to the left vocal fold. From the area function, the overall fundamental frequency of phonation decreases from 138 Hz ( $Q = 1$ ) to 113 Hz. The imbalance in tissue properties induces a phase offset between the lateral displacement waveforms of the left and right vocal folds, with the left vocal fold leading the right vocal fold in phase. Thus, with  $Q = 0.77$ , the following measures are computed: PA = 29.7%, AA = 0.2%, CIQ = 32.8%, and TL\* = -7.0 dB/octave.

#### 2. Effects of left–right phase asymmetry

Figure 7 displays plots of four salient glottal-related measures versus PA. Linear relationships exist between PA

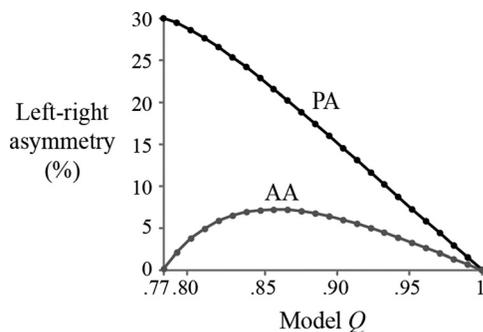


FIG. 5. Effects of changing model parameter  $Q$ , in increments of 0.01, on left–right phase asymmetry PA (black) and left–right amplitude asymmetry AA (gray).

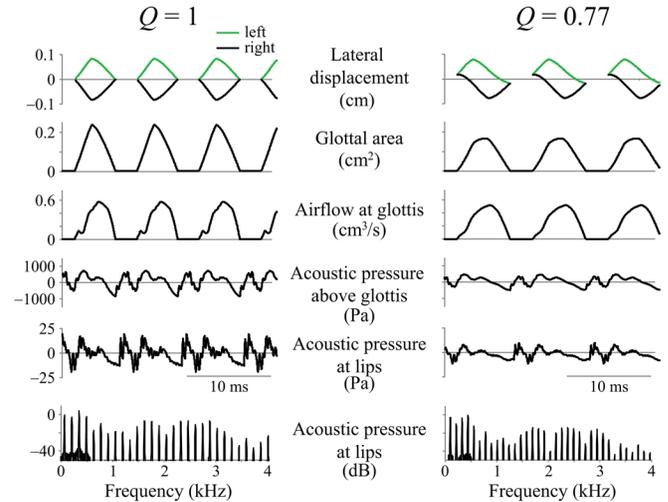


FIG. 6. (Color online) Waveforms from model simulations with  $Q = 1$  (left column) and  $Q = 0.77$  (right column).

and both PQ [Fig. 7(A)] and AS [Fig. 7(B)]. The corresponding changes in CIQ exhibit two patterns [Fig. 7(C)]. For PA less than 11%, there is an inverse linear trend with CIQ. Larger PA values do not have as significant an effect on CIQ. This relationship indicates progressively more abrupt closure during simulations with PA increasing from 0% to 11%. Figure 7(D) shows an overall inverse relationship between PA and spectral tilt measure TL\*. Therefore, vibratory phase asymmetry seems to be linked monotonically to spectral tilt. Although not plotted, there is an overall positive correlation between PA and the other acoustic measures ( $H1^* - H2^*$ ,  $H1^* - A1^*$ , and  $H1^* - A3^*$ ). These relationships are explored further in the subject data.

### B. Human subject data

#### 1. Prevalence of vocal fold vibratory asymmetry

Table I displays summary statistics of the HSV-based measures and acoustic measures separately for subjects with normal and disordered voices. Vocal fold vibration exhibits

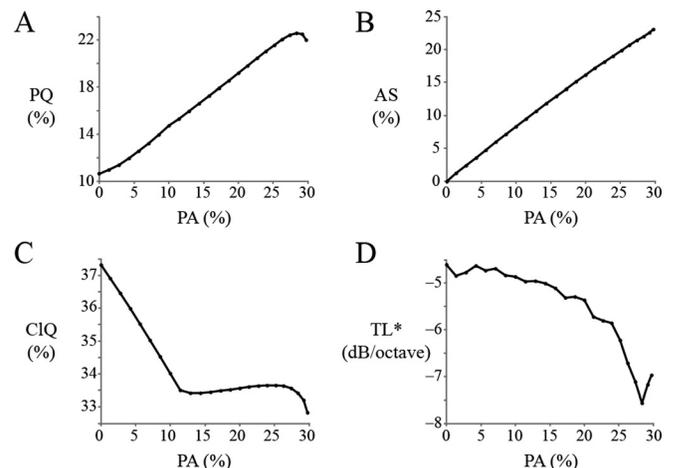


FIG. 7. Model-based covariations of left–right phase asymmetry PA with (A) plateau quotient PQ, (B) axis shift during closure AS, (C) closing quotient CIQ, and (D) spectral tilt TL\*.

TABLE I. Summary statistics of HSV-based asymmetry measures, HSV-based glottal area measures, and spectral measures of the radiated acoustic pressure waveform for human subject recordings<sup>a</sup>.

Measure	Units	Min	Max	$\mu$	$\sigma$
PA <sup>b</sup>	%	0, 1	12, 56	7, 20	5, 14
AA <sup>b</sup>	%	0, 0	16, 40	5, 11	5, 10
AS <sup>b</sup>	%	1, 1	11, 50	5, 16	4, 13
OQ	%	36, 38	85, 90	57, 66	18, 15
SQ	%	74, 51	224, 301	115, 134	52, 49
CIQ	%	17, 11	39, 55	26, 28	8, 9
PQ	%	10, 8	17, 29	14, 15	3, 4
H1*-H2*	dB	-1.7, -41.5	14.6, 21.4	3.6, 7.0	5.3, 8.5
H1*-A1*	dB	-5.6, -12.9	24.4, 27.3	12.3, 13.3	6.9, 7.2
H1*-A3*	dB	11.5, -8.4	28.7, 40.3	16.7, 20.2	5.6, 9.1
TL*	dB/octave	-12.5, -14.8	-5.8, 0.82	-7.2, -8.9	2.4, 3.2

<sup>a</sup>Minimum (Min), maximum (Max), average ( $\mu$ ), and standard deviation ( $\sigma$ ) are given for each measure across the subjects with no history of vocal pathology (first value,  $N=7$ ) and subjects exhibiting voice disorders (second value,  $N=46$ ).

<sup>b</sup>Directionality of the measure is removed by computing its magnitude.

complete glottal closure, with the largest degrees of PA and AS at  $\sim 50\%$ . This observation is in contrast to the model implementation of varying  $Q$ , in which controlled levels of PA are only obtainable up to 30% (Fig. 5).

## 2. Case studies

Figure 8 displays a snapshot of HSV and acoustic descriptors for a male adult subject with no history of voice disorders. The acoustic fundamental frequency is 136 Hz. PA measured from the medial digital kymogram is 6%. AS is  $-2\%$ , and AA is 0%. The spectral tilt TL\* of the acoustic magnitude spectrum is  $-6.4$  dB/octave. Glottal area waveform measures computed: OQ = 56%, SQ = 86%, CIQ = 29%, and PQ = 12%.

In contrast, Fig. 9 presents data from a male adult subject who underwent surgical treatment for cancer of the right vocal fold. Due to the surgical procedure, the tissue on the right vocal fold exhibits a reduction in mass and an increase in stiffness due to scar tissue and loss of superficial lamina

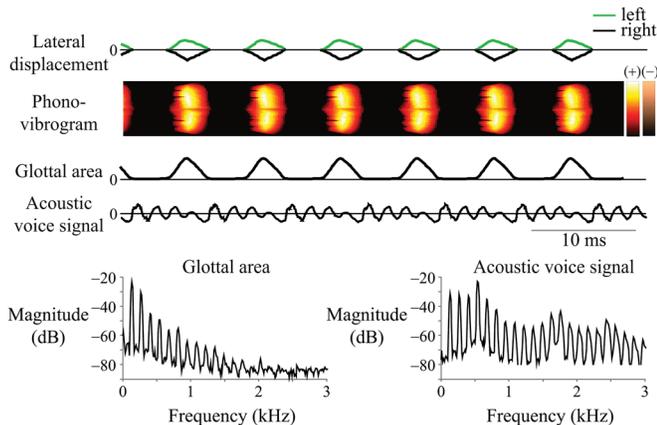


FIG. 8. (Color online) Case study of subject N1 with a low degree of left-right phase asymmetry PA (6%). Plots display the lateral displacement waveforms, phonovibrogram, glottal area waveform, radiated acoustic pressure waveform, and magnitude frequency spectra of the glottal area and radiated acoustic pressure.

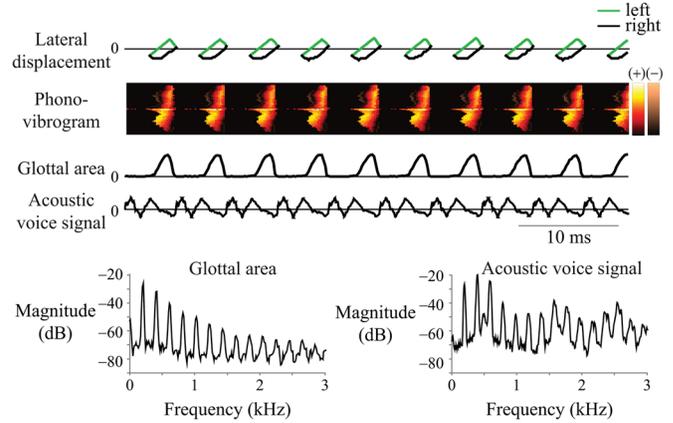


FIG. 9. (Color online) Case study of subject P13 with a high degree of left-right phase asymmetry PA ( $-51\%$ ). Cf. Fig. 8.

propria. The left-right imbalance produces asymmetric vocal fold vibration. The acoustic fundamental frequency is 197 Hz. Agreeing with model-based predictions, the stiffer (right) vocal fold leads the opposite vocal fold in phase, yielding a negative value for PA ( $-51\%$ ) and an OQ of 46%. The relative amplitude of the right vocal fold is reduced (AA = 11%). SQ and CIQ are 300% and 11%, respectively, signifying more rapid closure during the closing phase relative to that observed in the normal case. The spectral tilt TL\* of the disordered voice is  $-6.3$  dB/octave, which is notably similar to that in the normal case.

## 3. Acoustic correlates of glottal characteristics

Table II displays the intercorrelation matrix for the HSV-based and acoustic measures calculated from the subject data. In particular, none of the measures of vibratory asymmetry, nor plateau quotient, correlates to a statistically significant degree with any of the acoustic spectral measures. Glottal area measures OQ and CIQ, however, exhibit significant correlations with acoustic measures, ranging in strength from  $|r| = 0.27$  to  $|r| = 0.48$ .

Figure 10 displays scatter plots for four statistically significant pairwise correlations, with model trajectories overlaid. Figure 10(A) displays a wide variation in AA accompanying different values of PA, which is modeled in a nonlinear manner. Figure 10(B) shows a linear relationship between PA and AS that is exhibited by the model.

Figure 10(C) displays a moderate inverse relationship between CIQ and PA, with the model matching this relationship at low values of PA. Finally, Fig. 10(D) indicates that CIQ correlates with TL\* to a moderate degree, where the model line is computed by mapping phase asymmetry onto closing quotient values to simulate variations in CIQ [Fig. 7(C)].

## IV. DISCUSSION

Left-right vibratory phase asymmetry is simulated in a computational voice production model that consists of an asymmetric two-mass vocal fold model and nonlinear source-filter interaction with subglottal and supraglottal tracts. Model-based predictions are compared with human subject results obtained from time-synchronized measures of

TABLE II. Pearson's correlation coefficient  $r$  for significant pairwise relationships among average values of HSV-based asymmetry measures, HSV-based glottal area measures, and spectral measures of the radiated acoustic pressure waveform from all subject recordings ( $N = 53$ )<sup>a</sup>.

	AA <sup>b</sup>	AS <sup>b</sup>	OQ	SQ	CIQ	PQ	$H1^*-H2^*$	$H1^*-A1^*$	$H1^*-A3^*$	TL*
PA <sup>b</sup>	-0.42	0.93	—	0.45	-0.39	—	—	—	—	—
AA <sup>b</sup>		-0.36	—	0.28	—	—	—	—	—	—
AS <sup>b</sup>			—	0.44	-0.41	—	—	—	—	—
OQ				—	0.78	-0.35	0.28	0.27	0.40	-0.31
SQ					-0.63	—	—	—	-0.28	—
CIQ						—	0.32	0.31	0.48	-0.41
PQ							—	—	—	—
$H1^*-H2^*$								0.69	0.58	-0.52
$H1^*-A1^*$									0.67	-0.65
$H1^*-A3^*$										-0.85

<sup>a</sup>Correlations shown are statistically significant at a 95% confidence level.

<sup>b</sup>Directionality is removed unless pairwise correlations are computed between two HSV-based asymmetry measures.

laryngeal HSV and the acoustic voice signal. Studying the extended two-mass model independently from subject recordings (versus fitting the model to subject data) provides the power to compare and contrast data analysis from the different perspectives.

Theoretical predictions of glottal dynamics and the glottal area-to-acoustic pressure transformation are implemented through system nonlinearities, including acoustic loading due to subglottal and supraglottal tracts, the activation of additional spring constants during collision, and nonlinear equations for the intraglottal pressures. Future model improvements could incorporate additional damping constants during collision (Ishizaka and Flanagan, 1972) and more realistic descriptions of vocal fold aerodynamics, tissue deformation, and incomplete glottal closure (Pelorson *et al.*, 1994; Zañartu, 2010).

### A. Characteristics of the mathematical model

The original use of the model was to analyze parameters creating bifurcations and irregular behavior (Steinecke and Herzog, 1995). In contrast, the current study selects the range of the asymmetry factor  $Q$  that maintains periodicity, glottal closure, and a linear relationship with a measure of left-right phase asymmetry PA. The model predicts the degree and polarity of the relationship between PA and the axis shift during closure AS in the subject data, explaining 87% of the variance [Fig. 10(B)]. This strong relationship corroborates previous results from a group of 52 normal speakers, in which changes in PA accounted for 74% of the variance in AS (Mehta *et al.*, 2011), and from qualitative links between these two measures in voice patients (Švec *et al.*, 2007).

Although the model sustains periodic oscillations with PA magnitudes up to 30%, some human subjects exhibit

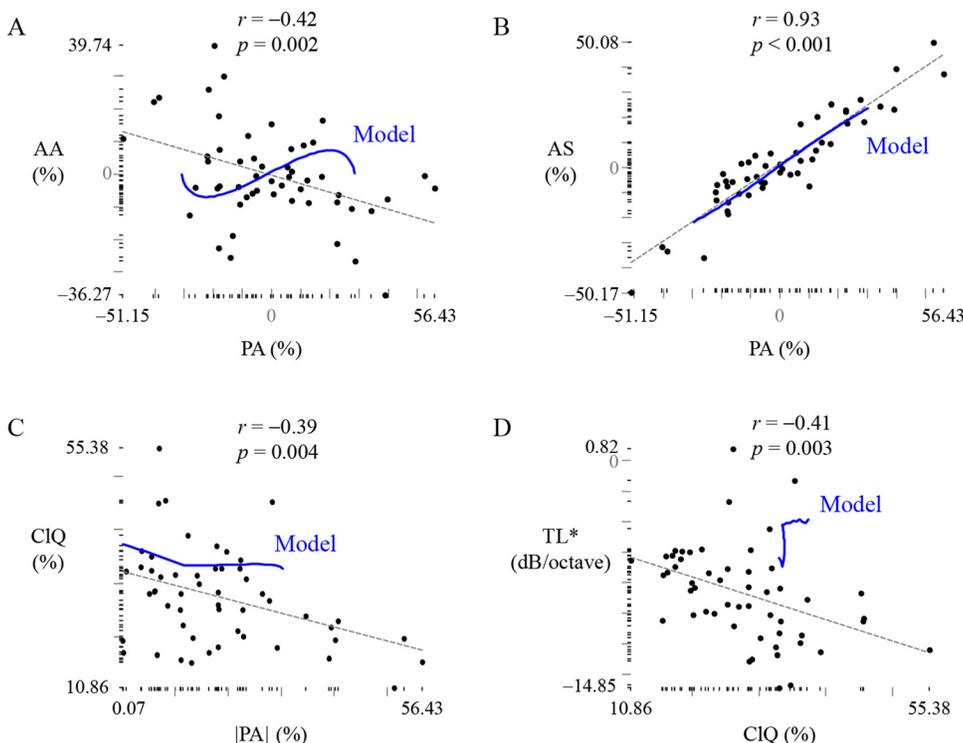


FIG. 10. (Color online) Scatter plots of statistically significant relationships in the subject data. (A) PA and AA, (B) PA and AS, (C)  $|PA|$  and CIQ, and (D) CIQ and TL\*. Pearson's  $r$  and  $p$ -value are indicated. Pairwise relationships obtained from the model simulations sweeping the  $Q$  parameter are overlaid on their respective plots. Note that the model trendline in (D) is derived from model data as CIQ is not an independent model parameter. Axes are linear, with tick marks positioned at the abscissa and ordinate of each point and longer tick marks (gray) at 10% or 2 dB/octave intervals.

levels of PA over 50% (Fig. 10). Future model modifications may take into account this observation, especially as it relates to subjects with vocal pathologies, and ongoing work is beginning to look at alternative implementations of asymmetric intraglottal flow (Erath *et al.*, 2011). In addition, the model assumption of a uniform rectangular glottis is violated in many of the subjects with voice disorders (exemplified in Fig. 9) whose glottal aperture is not rectangular in reality.

The parametric changes of  $Q$  in the model yield results that differ from pairwise relationships in the subject data between PA and physiological measures of left–right amplitude asymmetry AA [Fig. 10(A)] and plateau quotient PQ [high correlation in Fig. 7(A) and nonexistent correlation in Table II]. The present results point to a need for improvements in the model to predict certain features of the glottal area waveform. Future work could investigate alternative methods of controlling asymmetry and the effects of different baseline model parameters.

Concomitant changes in AA are not controlled for in the model (see Fig. 5) when obtaining the relationship between PA and measures quantifying the glottal area waveform and acoustic voice spectrum. Thus, changes in  $Q$  might not adequately represent voices with the pathologies included in this study. Also, cycle-to-cycle perturbations of the glottal area waveform are not modeled. Glottal correlates of acoustic shimmer and jitter have proven difficult to investigate, possibly owing to the sensitivity of the measures to turbulent noise in the signal and to the lack of understanding of exactly how to incorporate shimmer, jitter, and turbulent noise in such a model (Jiang and Zhang, 2002).

The model does not reflect the absolute value of CIQ exhibited by the subjects [Fig. 10(C)]. One potential cause is the use of a standard configuration of the model to simulate modal phonation that would be exhibited by one subject. In the current work, relative changes in asymmetry are of interest rather than absolute values. Ongoing efforts seek to extend vocal fold models to represent “virtual human subjects,” where multiple conditions and parametric variations are explored to statistically describe an entire population (Cook *et al.*, 2009).

Statements regarding effects of a single type of vibratory asymmetry cannot be made due to the complex nature of vocal fold vibration. Prior to the present investigation, it was unknown how the extended two-mass model would react to changes in left–right phase asymmetry. Models tend to exaggerate the effects of particular vocal fold kinematic conditions, and the current work places the model outputs side-by-side with data from human subjects to allow the reader to draw conclusions about the validity of the model.

## B. Physiological behaviors observed

Asymmetric vocal fold vibration has been associated with less abrupt glottal closure and, thus, less efficient sound production (Khosla *et al.*, 2008; Murugappan *et al.*, 2009). A direct consequence of this association is a presumed steeper spectral tilt of the source harmonics for higher values of PA. The model results of the current study concur with this hypothesis, where increases in PA correspond to more negative

values of  $TL^*$ , or steeper spectral tilt [Fig. 7(D)]. In the subject data, however, none of the measures of vocal fold vibratory asymmetry correlates directly with any of the acoustic spectral measures.

PA exhibits mild correlations with glottal area waveform measures SQ ( $r=0.45$ ,  $p=0.001$ ) and CIQ ( $r=-0.39$ ,  $p=0.004$ ) (Table II). This observation warrants further empirical studies to gain insight into the mechanisms that affect the closing phase of glottal airflow, which is theoretically more closely tied to acoustic sound production than glottal area measures. Nevertheless, statistically significant correlations are still obtained with glottal area–based measures of closing quotient, suggesting that some of the variance in the properties of the acoustic voice signal can be explained by vocal fold kinematics. In addition, whereas spectral tilt has been strongly linked to closing phase characteristics of the glottal airflow, acoustic measures such as  $H1^*-H2^*$  have been shown to depend on multiple source properties such as open quotient and glottal pulse skewing (Doval and d’Alessandro, 2006).

Airflow measurements potentially could yield additional information related to glottal pulse skewing that complements HSV-based measures of asymmetry and glottal area. Efforts are under way to add aerodynamic assessment to the HSV setup to gain insight into these relationships (Zañartu *et al.*, 2011). It is known that the glottal pulse shape varies depending on whether it is observed from the glottal area waveform or the glottal airflow waveform (Titze, 1984). Vocal tract inertance, respiratory system anatomy, and nonlinear source–filter interactions influence the glottal airflow waveform and thus the energy propagation of airflow/pressure distributions to the lips.

An expected, although weak, inverse correlation ( $r=-0.41$ ,  $p=.003$ ) was found between spectral tilt  $TL^*$  and closing quotient CIQ in the subject sample [Fig. 10(D)]. Thus, some voices with higher values of CIQ (less abrupt closure) produce phonation with more negative values of  $TL^*$  (steeper rolloff). Titze (2006) proposed that the closing phase duration is only an indirect measure of maximum area declination rate, which, in turn, influences maximum flow declination rate, sound pressure level, and spectral tilt. Thus, a low correlation between CIQ and  $TL^*$  is not completely unexpected. The model displays atypical behavior when CIQ is plotted against  $TL^*$  partly because the closing quotient is not the independent variable in this study.

## C. Potential clinical application

Results of the present study provide additional support for the clinical application of high-speed imaging as a way to increase the accuracy with which salient deficits in vocal fold function can be identified and pinpointed to complement the current reliance on videostroboscopic assessment (Deliy-ski *et al.*, 2008; Bonilha *et al.*, 2011). The findings demonstrate the importance of gaining a better understanding of the relationships between vocal fold vibratory properties and characteristics of the sound that is produced. Clinicians can then focus assessment and treatment efforts on aspects of the phonatory process that have the most impact on voice function and quality.

The actual clinical utility of the current results, however, is clearly limited by the relatively low correlations that were obtained, providing motivation for continued efforts to account for more of the unexplained variance in the acoustic and HSV-based measures. Such efforts could include the addition of aerodynamic measures and methods for capturing the three-dimensional motion of the vocal folds to more comprehensively describe the complex fluid–structure–acoustic interaction that takes place during phonation.

Considerable physiological variations exist across subjects, regardless of vocal health, and future protocols would benefit from collecting multiple measurements of the same speaker at various loudness and pitch gestures. A higher-pitch condition was available in a small subset of subjects in the current study. Preliminary analysis of this subset, although, did not reveal any additional information regarding relationships among vocal fold vibratory phase asymmetry, glottal area measures, and acoustic spectral measures.

## V. CONCLUSION

Laryngeal high-speed videoendoscopy recordings of human subjects reveal that there is no direct correlation between vocal fold vibratory asymmetries and acoustic spectral tilt measures. This result calls for the development of improved acoustic measures and investigations into the effect of asymmetric vocal fold vibration on glottal airflow and the associated impact on voice source properties and vocal efficiency. Phase asymmetry between the lateral displacements of the left and right vocal folds plays only one role in shaping the modulation of the glottal airflow and, ultimately, the spectral characteristics of the acoustic voice signal.

## ACKNOWLEDGMENTS

The work of T.F.Q. was supported by the Department of Defense under Air Force Contract No. FA8721-05-C-0002. The opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government. The work of the other authors was supported by grants from the NIH National Institute on Deafness and Other Communication Disorders (T32 DC00038 and R01 DC007640) and by the Institute of Laryngology and Voice Restoration. The paper's contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Berry, D. A., Herzel, H., Titze, I. R., and Story, B. H. (1996). "Bifurcations in excised larynx experiments," *J. Voice* **10**, 129–138.

Boersma, P., and Weenink, D. (2009). "Praat: Doing phonetics by computer (version 5.1.40) [computer program]," <http://www.praat.org> (Last viewed July 13, 2009).

Bonilha, H. S., Deliyski, D. D., and Gerlach, T. T. (2008). "Phase asymmetries in normophonic speakers: Visual judgments and objective findings," *Am. J. Speech Lang. Pathol.* **17**, 367–376.

Bonilha, H. S., Gerlach, T. T., Whiteside, J. P., and Deliyski, D. D. (2011). "Vocal fold phase asymmetries in patients with voice disorders: A study across visualization techniques," *Am. J. Speech Lang. Pathol.* (in press).

Cook, D. D., Nauman, E., and Mongeau, L. (2009). "Ranking vocal fold model parameters by their influence on modal frequencies," *J. Acoust. Soc. Am.* **126**, 2002–2010.

Deliyski, D. D. (2005). "Endoscope motion compensation for laryngeal high-speed videoendoscopy," *J. Voice* **19**, 485–496.

Deliyski, D. D., and Petrushev, P. (2003). "Methods for objective assessment of high-speed videoendoscopy," *Proceedings of the International Conference on Advances in Quantitative Laryngology*, Vol. 1(28), pp. 1–16.

Deliyski, D. D., Petrushev, P. P., Bonilha, H. S., Gerlach, T. T., Martin-Harris, B., and Hillman, R. E. (2008). "Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution," *Folia Phoniatr. Logop.* **60**, 33–44.

Dormand, J. R., and Prince, P. J. (1980). "A family of embedded Runge-Kutta formulae," *J. Comput. Appl. Math.* **6**, 19–26.

Doval, B., and d'Alessandro, C. (2006). "The spectrum of glottal flow models," *Acta Acust. Acust.* **92**, 1026–1046.

Erath, B. D., Peterson, S. D., Zañartu, M., Wodicka, G. R., and Plesniak, M. W. (2011). "A theoretical model of the pressure field arising from asymmetric intraglottal flows applied to a two-mass model of the vocal folds," *J. Acoust. Soc. Am.* **130**, 389–403.

Fant, G. (1997). "The voice source in connected speech," *Speech Commun.* **22**, 125–139.

Haben, C. M., Kost, K., and Papagiannis, G. (2003). "Lateral phase mucosal wave asymmetries in the clinical voice laboratory," *J. Voice* **17**, 3–11.

Hanson, H. M., and Chuang, E. S. (1999). "Glottal characteristics of male speakers: Acoustic correlates and comparison with female data," *J. Acoust. Soc. Am.* **106**, 1064–1077.

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., Perkell, J. S., Guiod, P. C., and Goldman, S. L. (1995). "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," *J. Speech Hear. Res.* **38**, 1212–1223.

Iseli, M., Shue, Y.-L., and Alwan, A. (2007). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," *J. Acoust. Soc. Am.* **121**, 2283–2295.

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.

Ishizaka, K., and Ishiki, N. (1976). "Computer simulation of pathological vocal-cord vibration," *J. Acoust. Soc. Am.* **60**, 1193–1198.

Jiang, J. J., and Zhang, Y. (2002). "Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds," *J. Acoust. Soc. Am.* **112**, 2127–2133.

Khosla, S., Murugappan, S., and Gutmark, E. (2008). "What can vortices tell us about vocal fold vibration and voice production," *Curr. Opin. Otolaryngol. Head Neck Surg.* **16**, 183–187.

Khosla, S., Murugappan, S., Paniello, R., Ying, J., and Gutmark, E. (2009). "Role of vortices in voice production: Normal versus asymmetric tension," *Laryngoscope* **119**, 216–221.

Lohscheller, J., Döllinger, M., McWhorter, A. J., and Kunduk, M. (2008a). "Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms," *Ann. Otol. Rhinol. Laryngol.* **117**, 484–493.

Lohscheller, J., Eysholdt, U., Toy, H., and Döllinger, M. (2008b). "Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics," *IEEE Trans. Med. Imaging* **27**, 300–309.

Maunsell, R., Ouaknine, M., Giovanni, A., and Crespo, A. (2006). "Vibratory pattern of vocal folds under tension asymmetry," *Otolaryngol. Head Neck Surg.* **135**, 438–444.

Mehta, D. D. (2010). "Impact of human vocal fold vibratory asymmetries on acoustic characteristics of sustained vowel phonation," *Doctoral dissertation, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA.*

Mehta, D. D., Deliyski, D. D., Quatieri, T. F., and Hillman, R. E. (2011). "Automated measurement of vocal fold vibratory asymmetry from high-speed videoendoscopy recordings," *J. Speech. Lang. Hear. Res.* **54**, 47–54.

Mehta, D. D., Deliyski, D. D., Zeitels, S. M., Quatieri, T. F., and Hillman, R. E. (2010). "Voice production mechanisms following phonosurgical treatment of early glottic cancer," *Ann. Otol. Rhinol. Laryngol.* **119**, 1–9.

Mehta, D. D., and Hillman, R. E. (2008). "Voice assessment: Updates on perceptual, acoustic, aerodynamic, and endoscopic imaging methods," *Curr. Opin. Otolaryngol. Head Neck Surg.* **16**, 211–215.

Murugappan, S., Khosla, S., Casper, K., Oren, L., and Gutmark, E. (2009). "Flow fields and acoustics in a unilateral scarred vocal fold model," *Ann. Otol. Rhinol. Laryngol.* **118**, 44–50.

- Niimi, S., and Miyaji, M. (2000). "Vocal fold vibration and voice quality," *Folia Phoniatr. Logop.* **52**, 32–38.
- Pelorson, X., Hirschberg, A., van Hassel, R. R., Wijnands, A. P. J., and Aurégan, Y. (1994). "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation: Application to a modified two-mass model," *J. Acoust. Soc. Am.* **96**, 3416–3431.
- Qiu, Q., Schutte, H. K., Gu, L., and Yu, Q. (2003). "An automatic method to quantify the vibration properties of human vocal folds via videokymography," *Folia Phoniatr. Logop.* **55**, 128–136.
- Schwarz, R., Hoppe, U., Schuster, M., Wurzbacher, T., Eysholdt, U., and Lohscheller, J. (2006). "Classification of unilateral vocal fold paralysis by endoscopic digital high-speed recordings and inversion of a biomechanical model," *IEEE Trans. Biomed. Eng.* **53**, 1099–1108.
- Steinecke, I., and Herzel, H. (1995). "Bifurcations in an asymmetric vocal-fold model," *J. Acoust. Soc. Am.* **97**, 1874–1884.
- Story, B. H., and Titze, I. R. (1995). "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.* **97**, 1249–1260.
- Švec, J. G., Šram, F., and Schutte, H. K. (2007). "Videokymography in voice disorders: What to look for?" *Ann. Otol. Rhinol. Laryngol.* **116**, 172–180.
- Takemoto, H., Honda, K., Masaki, S., Shimada, Y., and Fujimoto, I. (2006). "Measurement of temporal changes in vocal tract area function from 3D cine-MRI data," *J. Acoust. Soc. Am.* **119**, 1037–1049.
- Tao, C., Zhang, Y., Hottinger, D. G., and Jiang, J. J. (2007). "Asymmetric airflow and vibration induced by the Coanda effect in a symmetric model of the vocal folds," *J. Acoust. Soc. Am.* **122**, 2270–2278.
- Titze, I., Riede, T., and Popolo, P. (2008). "Nonlinear source-filter coupling in phonation: Vocal exercises," *J. Acoust. Soc. Am.* **123**, 1902–1915.
- Titze, I. R. (1984). "Parameterization of the glottal area, glottal flow, and vocal fold contact area," *J. Acoust. Soc. Am.* **75**, 570–580.
- Titze, I. R. (2002). "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," *J. Acoust. Soc. Am.* **111**, 367–376.
- Titze, I. R. (2006). "Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation," *J. Speech. Lang. Hear. Res.* **49**, 439–447.
- Titze, I. R. (2008). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**, 2733–2749.
- van den Berg, J. (1958). "Myoelastic-aerodynamic theory of voice production," *J. Speech Hear. Res.* **1**, 227–244.
- Verdonck-de Leeuw, I. M., Festen, J. M., and Mahieu, H. F. (2001). "Deviant vocal fold vibration as observed during videokymography: The effect on voice quality," *J. Voice* **15**, 313–322.
- Weibel, E. R. (1963). *Morphometry of the Human Lung*, 1st ed. (Springer, New York), p. 139.
- Wurzbacher, T., Schwarz, R., Döllinger, M., Hoppe, U., Eysholdt, U., and Lohscheller, J. (2006). "Model-based classification of nonstationary vocal fold vibrations," *J. Acoust. Soc. Am.* **120**, 1012–1027.
- Xue, Q., Mittal, R., Zheng, X., and Bielamowicz, S. (2010). "A computational study of the effect of vocal-fold asymmetry on phonation," *J. Acoust. Soc. Am.* **128**, 818–827.
- Zañartu, M. (2010). "Acoustic coupling in phonation and its effect on inverse filtering of oral airflow and neck surface acceleration," Doctoral dissertation, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.
- Zañartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., and Hillman, R. E. (2011). "Observation and analysis of *in vivo* vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study," *J. Acoust. Soc. Am.* **129**, 326–339.
- Zañartu, M., Mongeau, L., and Wodicka, G. R. (2007). "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.* **121**, 1119–1129.
- Zeitels, S. M., Blitzer, A., Hillman, R. E., and Anderson, R. R. (2007). "Foresight in laryngology and laryngeal surgery: A 2020 vision," *Ann. Otol. Rhinol. Laryngol. Suppl.* **116**, 1–16.
- Zhang, Y., Tao, C., and Jiang, J. J. (2006). "Parameter estimation of an asymmetric vocal-fold system from glottal area time series using chaos synchronization," *Chaos* **16**, 023118.