Investigating the relationship between glottal area waveform shape and harmonic magnitudes through computational modeling and laryngeal high-speed videoendoscopy

Gang Chen[1], *Robin A. Samlan*[2], *Jody Kreiman*[2], *Abeer Alwan*[1]

[1]Department of Electrical Engineering, University of California, Los Angeles, USA
[2]Department of Head and Neck Surgery, School of Medicine, University of California, Los Angeles, USA
{gangchen,alwan}@ee.ucla.edu rsamlan@mednet.ucla.edu jkreiman@ucla.edu

Abstract

The glottal open quotient (OQ) is often associated with the amplitude of the first source harmonic relative to the second (H1*-H2*), which is assumed to be one cause of a change in vocal quality along a breathy-to-pressed continuum. The association between OQ and H1*-H2* was investigated in a group of 5 human subjects and also in a computational voice production simulation. The simulation incorporated a parametric voice source model into a nonlinear source-filter framework. H1*-H2* and OQ were measured synchronously from audio recordings and high-speed laryngeal videoendoscopy of "glide" phonations in which quality varied continuously from breathy to pressed. Analyses of individual speakers showed large differences in the relationship between OQ and H1*-H2*. The variability in laryngeal high-speed data was consistent with simulation results, which showed that the relationship between OQ and H1*-H2* depended on mean glottal area, a parameter associated with the degree of source-filter interaction and not directly measurable from high-speed video of the vocal folds. In addition, H1*-H2* may change with increasing glottal gap size; this change contributes to the observed variability in the relationship between H1*-H2* and OQ.

Index Terms: harmonic magnitudes, laryngeal high-speed videoendoscopy, glottal area waveform, open quotient

1. Introduction

Many studies have investigated the acoustic and perceptual consequences of changes in glottal pulse shapes. Increases in open quotient (OQ, the relative amount of time the glottis is open within a glottal vibratory cycle) are widely assumed to be the physical precursors of perceived breathiness, in part because of consequent changes in the relative amplitudes of the first two harmonics of the voice source (denoted H1*-H2*, when harmonic amplitudes are measured from the audio signal recorded at the mouth and then corrected for the effects of vocal tract resonances [1, 2]). When OQ increases, the length of glottal open phase becomes closer to the glottal cycle, leading to a stronger fundamental component in the source spectrum (assuming all other influences, including pulse skewness, are constant) [3]. This increase in the fundamental component results in increased H1*-H2*, which is a primary presumptive cause of the change in vocal quality [4].

Most empirical studies used electroglottographic (EGG) data or inverse-filtered acoustic signals, and varying levels of correlation between H1*-H2* and OQ have been reported. In [5], H1*-H2* estimated from acoustic spectra was modestly correlated with the adduction quotient (defined as vocal fold contact time/period, or 1-OQ) measured from EGG and airflow data for 20 female speakers (r=-0.46 for EGG measures, and r=-0.69 for airflow data). In [6], four tokens of the vowel /a/ recorded from seven speakers were inverse-filtered and fitted with a Liljencrants-Fant (LF) source model [7]. OQ was then measured from the best-fitting model. Results showed that H1*-H2* and OQ were positively correlated for 17/28 tokens, negatively correlated for 4/28 tokens, and uncorrelated for 7/28 tokens. In recent work [8], H1*-H2*, OQ, and glottal area waveform skewness were measured synchronously from audio recordings and high-speed video images of the larynges of six speakers. Results showed that H1*-H2* could be predicted by OQ, glottal area waveform skewness, and fundamental frequency (F0) with good accuracy (together accounting for an average of 74% of the variance), but the regression model parameters were speaker dependent. Many other studies (e.g., [3, 7, 9, 10]) have examined the relationship between harmonic amplitudes and glottal configuration in the context of models of the voice source within a linear source-filter speech production framework [11]. For example, in the LF model the relationship is expressed as H1*-H2*=-6+0.27exp(0.055 OQ) [3].

However, in [12], analyses of singing (and other) voice showed that the LF model is unable to match some measured H1*-H2* values. In that study, it was suggested that either the LF model is limited, or else source-filter interaction is large enough to hinder the linear source-filter approximation in some phonatory modes. Nonlinear source-filter coupling between subglottal and supraglottal tracts has been shown to significantly affect source properties [13, 14, 15, 16]. This coupling has the effect of controlling the well-documented pulse skewing of the glottal airflow waveform, which is critical for determining spectral tilt (e.g., H1*-H2*) [12]. A recent study [17] performed simulation using a computational kinematic model of the vocal folds [18] and incorporated the interactions between vocal fold vibration and the trachea/vocal tract airway systems [14, 19, 20]. Simulation results showed that H1*-H2* sometimes increased and sometimes decreased with increased separation of the vocal processes (OQ values were not reported). In [21], analyses of high-speed recordings of the vocal folds showed that H1*-H2* increased with increasing glottal gap size when glottal closure was incomplete. Although different glottal area and flow signals were simulated by varying control parameters such as degree of vocal fold adduction, surface bulging, vibratory nodal point, and supraglottal constriction, the kinematic model in [17] does not offer simple control over OQ or glottal gap size.

In summary, despite the insights that modeling studies have provided regarding the relationship between OQ and H1*-H2*, we cannot currently account for all reported observations of natural data. Although experimental studies suggest that the relationship between H1*-H2* and OQ may be variable, modeling studies do not currently fully explain the observed variability. This study used natural data to examine the relationship between H1*-H2* (measured from recorded acoustic signals), OQ, and glottal gap size (measured synchronously from high-speed video images of the vibrating vocal folds). Simulations of the glottal area waveforms were then performed using a parametric model of the voice source [22] that offers direct control of the glottal pulse shape, paired with nonlinear source-filter interactions to simulate glottal flow and the radiated acoustic signals. The effect of absolute glottal area was investigated and model simulation results were compared to those from laryngeal high-speed recordings of human subjects.

2. Data and methods

2.1. Human subject data

2.1.1. Subject selection and data acquisition

Synchronous audio recordings and high-speed videoendoscopic images of the vocal folds were collected from five phonetically knowledgeable subjects, four males (speakers 1-4) and one female (speaker 5). These speakers were asked to sustain the vowel /i/ and gradually change their phonations from breathy to pressed while holding F0 and vowel quality as steady as possible. The vowel /i/ was selected to optimize the view of the vocal folds [23]; across tokens vowel quality ranged from /I/ to approximately cardinal vowel /ɛ/. High-speed images of the vocal folds were recorded using a Phantom V210 camera (Vision Research, Wayne, NJ) at a sampling rate of 10,000 frames/s, with a resolution of 208×352 pixels. The camera was mounted on a Glidecam Camcrane 200 (Glidecam Industries, Kingston, MA). Audio signals were synchronously recorded with a Brüel & Kjær microphone (1.27 cm diameter; type 4193-L-004) and directly digitized at a sampling rate of 60 kHz. The audio recordings were later downsampled to 16 kHz for analysis. Synchronized audio and high-speed images were recorded for 6 s.

2.1.2. Measures from high-speed imaging

Glottal area waveforms of the complete utterances were extracted using "GlotAnTools", a software toolkit that automatically segments the glottal area from high-speed images [24]. Following [8, 25], each cycle of glottal vibration was tracked from the extracted glottal area waveforms by marking the first instants of glottal opening when glottal closure was complete. When no complete glottal closure occurred, the moments of minimal glottal area were tracked. For each individual cycle of phonation, OQ was calculated as the time from the first opening instant to the onset of maximum closure (or minimum area), divided by the length of the current glottal cycle. DC (i.e., the glottal gap size) was defined as the minimum glottal area normalized by the maximum glottal area in each glottal cycle. OQ and DC values were smoothed over 100 ms windows.

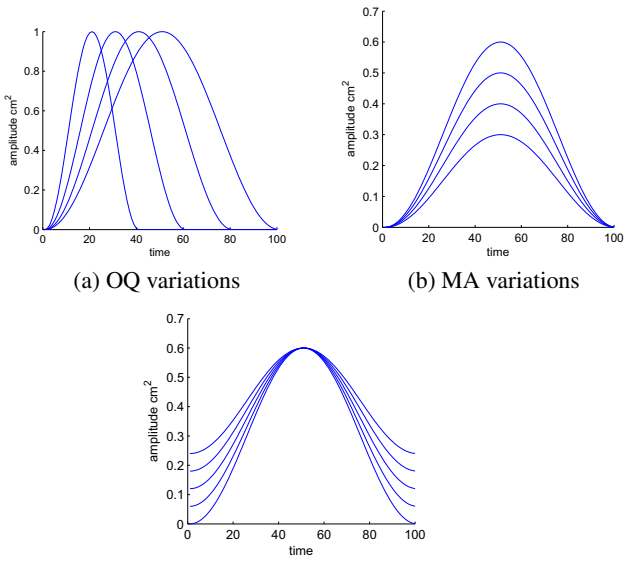
2.1.3. Acoustic measures

H1*-H2* was measured pitch-synchronously from the audio signals with VoiceSauce software [26] using an analysis window of eight periods with a 1 ms shift. F0 values were obtained from the STRAIGHT algorithm [27] to determine the period of a glottal cycle. Values were aligned with glottal area waveforms extracted from the imaging signal for subsequent analysis. The harmonic magnitudes, H1* and H2*, were calculated from the speech spectrum and corrected for the effects of the first two formant frequencies using the formula in [2]. The formant frequencies were estimated using the "Snack Sound Toolkit" software [28].

2.2. Computational model simulation

2.2.1. Generating glottal area waveforms

The parametric voice source model in [22] (denoted EE2) was chosen for this study to allow for direct control of the glottal area pulse shape. The EE2 model is a modified version of the EE model [29], which has been shown to be more effective in capturing the observed glottal **area** waveforms than the LF model [7, 29]. To generate glottal area waveforms, in all simulations the EE2 model parameters "asymmetry coefficient", "speed of opening", and "speed of closing" were set to a constant value of 0.5 (see [22] for definition of model parameters). The parameter OQ was varied from 0.3 to 1 with a step size of 0.05. Figure 1a shows sample waveforms obtained by varying OQ. The maximum amplitude (MA) of the glottal area waveform was varied from 0.1 cm^2 to 0.6 cm^2 with a step size of 0.05 cm^2, by multiplying the original waveform (Figure 1a) with different MA values (scaling factors) (Figure 1b). The range of MA was chosen according to [18]. Varying MA provides a way to simulate the relationship between H1*-H2* and OQ under varying degrees of source-filter interaction, because the degree of interaction depends on the mean glottal area [14]. In addition, although the source waveforms in Figure 1b have the same OQ and skewness (differing only by a scaling factor), the maximum area declination rates (MADR) vary across waveforms. The MADR is directly related to the maximum flow declination rates (MFDR), an important quantity highly correlated with vocal intensity [30]. DC (i.e., the glottal gap size, or minimum glottal area) was varied from 0 to 0.4 with a step size of 0.05, as demonstrated in Figure 1c. The glottal gaps observed in the high-speed images in this study extended through the cartilaginous glottis into some or all of the membranous glottis. OQ was set to 1 when varying the DC offset in generating the glottal area waveforms.



(a) OQ variations (b) MA variations

(c) DC variations

Figure 1: *Generated glottal area waveforms using the EE2 source model.*

2.2.2. *Simulating nonlinear source-filter interactions*

The glottal area was acoustically coupled to the trachea and vocal tract airway system [14, 19, 31]. The resulting glottal flow was determined by the interaction of the glottal area with the time-varying acoustic pressures present just inferior and superior to the glottis. The effect of turbulence was approximated by adding a noise component to the glottal flow signal using the method in [17]. These nonlinear source-filter interactions were simulated using the software toolkit "LeTalker" [32], a Matlab program to simulate human speech production. The parametric voice source model EE2, together with LeTalker, allowed for observation of the simulated glottal area, glottal flow, and radiated acoustic signals. H1*-H2* was measured from simulated glottal flow signals using custom Matlab code, following the settings described in Section 2.1.3.

3. Results

3.1. Human subject data

Figure 2 shows H1*-H2* and OQ for glide phonations from speakers 1-5. For speakers 1 and 5, the relationship between H1*-H2* and OQ is positive and approximately linear, despite the fluctuation of H1*-H2* when OQ is close to 1 (discussed below). For speaker 3, H1*-H2* is negatively correlated with OQ when OQ is below 0.75 and H1*-H2* is above 2 dB. This is not unexpected, because similar cases have been reported by previous studies. For example, for 3 out of 6 speakers in [8], OQ showed a negative regression coefficient in predicting H1*-H2* when OQ was below a certain threshold (thresholds were manually selected and ranged from 0.65 to 0.8). For speakers 2 and 4, H1*-H2* remains almost constant as OQ increases from 0.5 to 1. These varying patterns suggest that the relationship between H1*-H2* and phonatory characteristics may be speaker dependent.

Figure 3 shows H1*-H2* and DC for the same phonations from speakers 1-5. For speakers 1 and 4, H1*-H2* increases with increasing DC, consistent with the findings in [1, 21]. For speakers 2 and 5, H1*-H2* is negatively correlated with DC. For speaker 3, H1*-H2* only slightly decreases with increasing DC when DC is above 0.15. Despite this variability, the cases in which glottal gaps were observed are typically assigned an OQ of 100% [33], or close to 100% ([8, 25], also in the current study). Thus, the variability in H1*-H2 with varying DC partially contributes to the observed variability in the relationship between H1*-H2* and OQ in previous studies

3.2. Model simulations

Figures 4 and 5 show the effects of OQ and MA on H1*-H2* in the model simulations. When MA is relatively small (e.g., 0.1 cm^2 as in Figure 5a), H1*-H2* increases monotonically with increasing OQ and the relationship is approximately linear, similar to the observations in Figures 2a and 2e. When the MA is relatively large (e.g., 0.6 cm^2 as in Figure 5b), H1*-H2* first increases and then slightly decreases with increasing OQ. This is partially attributed to an increased degree of source-filter interaction. When OQ increases, the mean glottal area also increases, leading to a higher degree of source-filter interaction [14]. Recall that this interaction has the effect of "skewing" the glottal flow waveform, which results in decreased H1*-H2* [12]. This result is consistent with that in [17], where H1*-H2* sometimes increased and sometimes decreased with increased vocal process separation.

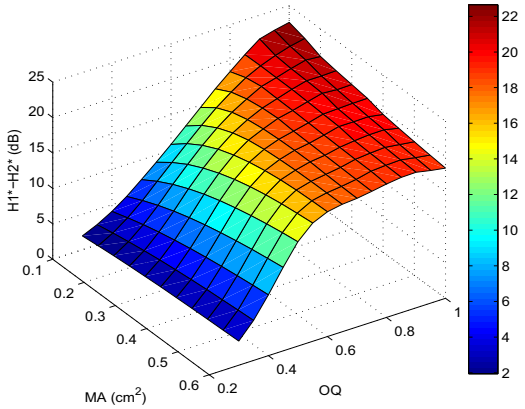
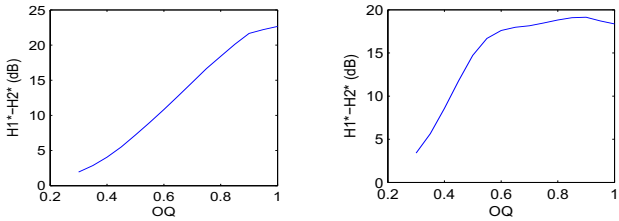


Figure 4: *Effects of changing OQ and MA, in increments of 0.05, on H1*-H2*. Color visualizes values of H1*-H2* on the z-axis.*



(a) MA=0.1 cm^2 (b) MA=0.6 cm^2

Figure 5: *Relationship between H1*-H2* and OQ for different MA values.*

Figures 6 and 7 show the effects of DC and MA on H1*-H2*. When MA is relatively large (e.g., 0.6 cm^2 as in Figure 7b), H1*-H2 decreases monotonically with increasing DC. This simulated case is similar to the human data in Figures 3b and 3e. Similar to the effect of increasing OQ, the increased DC also results in increased mean glottal area, which leads to a higher degree of source-filter interaction. The increased glottal flow waveform skewness caused by this interaction might have contributed to decreased H1*-H2*. When MA is relatively small (e.g., 0.1 cm^2 as in Figure 7a), H1*-H2* varies very little with increasing DC (about 2 dB).

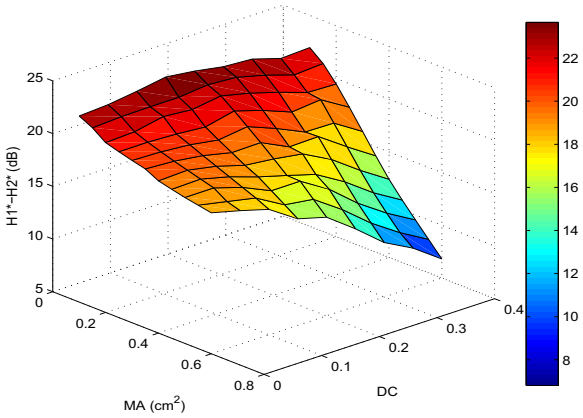
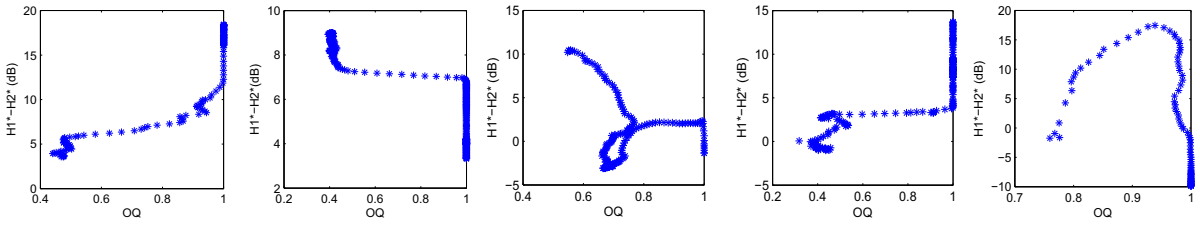
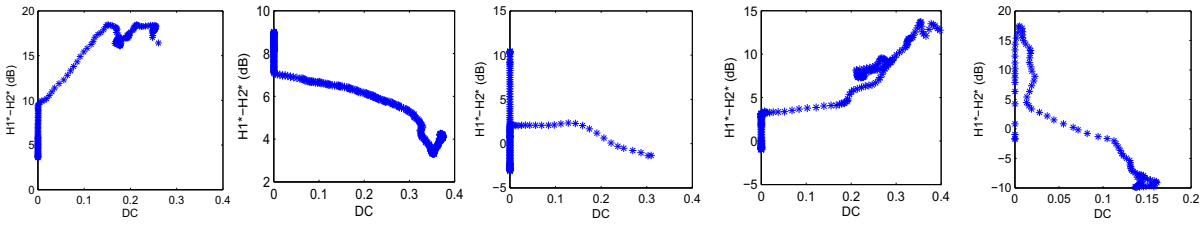


Figure 6: *Effects of changing DC and MA, in increments of 0.05, on H1*-H2*. Color visualizes values of H1*-H2* on the z-axis.*



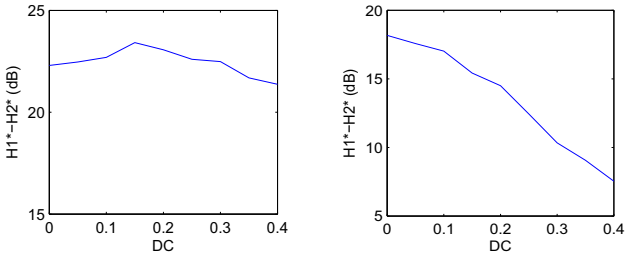
(a) Speaker 1 (b) Speaker 2 (c) Speaker 3 (d) Speaker 4 (e) Speaker 5

Figure 2: *H1*-H2* vs. OQ for speakers 1-5.*



(a) Speaker 1 (b) Speaker 2 (c) Speaker 3 (d) Speaker 4 (e) Speaker 5

Figure 3: *H1*-H2* vs. DC for speakers 1-5.*



(a) MA=0.1 cm^2 (b) MA=0.6 cm^2

Figure 7: *Relationship between H1*-H2* and DC for different MA values.*

4. Discussion

Standard laryngoscopy allows only a determination of the relative size of anatomical structures. Due to varying distances between the laryngoscope and the glottis across recordings, laryngoscopic recordings performed in separate sessions are not comparable. While methods have been proposed to measure the absolute scale of anatomical structures (e.g., using a laser projection device [34] or a stereo-endoscopy system [35]), they have not been widely applied. Although model simulations showed that the relationship between H1*-H2* and OQ depends on the maximum amplitude of the glottal area waveform, the absolute glottal area (in cm^2) is not directly measurable from high-speed images of the vocal folds. Previous studies (e.g., [8, 36]) based on high-speed laryngoscopy have mostly relied on time domain measures (e.g., open quotient, speed quotient, closing quotient, and pulse skewness). Even though time resolution could increase with increasing recording frame rate, it is still possible that the result derived from laryngeal high-speed recordings could be somewhat incomplete or inconclusive, especially when the acoustic variability associated with absolute glottal area is strong due to source-filter interactions. The simulation results in this study may also provide a possible explanation for the large interspeaker variability and weak correlations between time domain measures and acoustic measures reported in previous high-speed laryngoscopy-based studies [8, 36].

Although a kinematic model (e.g., [18]) might have provided more physiologically-realistic glottal area waveforms, such a model was not used in this study because it does not provide direct control over OQ and glottal gap size. In natural data, the maximum glottal area within each glottal cycle might also change when a speaker changes OQ [37]. Therefore, the simulations in this study, which varied source parameters while fixing maximum glottal area, are only approximations. Nevertheless, the simulation results may be interpreted as showing that the resultant H1*-H2* could be affected by absolute glottal area, even if all the time quotient measures are the same. Under the nonlinear source-filter framework, the absolute glottal area could be considered to provide an additional degree of variability in the acoustic parameters.

5. Conclusion

This study investigated the relationship between H1*-H2*, OQ, and glottal gap size. Analyses of synchronous audio and laryngeal high-speed video recordings showed that the effects of OQ and glottal gap size on H1*-H2* may be variable and speaker dependent. Model simulations supported the observed variabilities and suggested that this relationship depends on mean glottal area, a parameter associated with the degree of source-filter interaction but not directly measurable from high-speed images of the vocal folds. H1*-H2* may increase or decrease with increasing glottal gap size, allowing more variability of relationship between H1*-H2* and OQ to be observed. Future work will include recording other vowels using a flexible endoscope, as well as quantifying the absolute glottal area to aid analyses.

6. Acknowledgements

This work was supported in part by NSF Grant No. IIS-1018863 and by NIH/NIDCD Grant Nos. DC01797 and DC011300. We thank Juergen Neubauer, Marc Garellek and Dinesh Chhetri for help recording high-speed images. We also thank Michael Döllinger and the Department of Phoniatrics and Pediatric Audiology of the University Hospital, Erlangen, Germany, for generously supplying the GlotAnTools software.

7. References

[1] H. M. Hanson, "Glottal characteristics of female speakers: Acoustic correlates," *J. Acoust. Soc. Am.*, vol. 101, pp. 466–481, 1997.

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

[3] G. Fant, "The LF-model revisited. Transformations and frequency domain analysis," *STL-QPSR*, vol. 36, pp. 119–156, 1995.

[4] D. Klatt and L. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.*, vol. 87, pp. 820–857, 1990.

[5] E. B. Holmberg, R. E. Hillman, J. S. Perkell, P. Guiod, and S. L. Goldman, "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," *J. Speech Hear. Res.*, vol. 38, pp. 1212–1223, 1995.

[6] M. Swerts and R. Veldhuis, "The effect of speech melody on voice quality," *Speech Commun.*, vol. 33, no. 4, pp. 297–303, 2001.

[7] G. Fant, J. Liljencrants, and Q. Lin, "A four-parameter model of glottal flow," *STL-QPSR*, vol. 4, pp. 1–13, 1985.

[8] J. Kreiman, Y.-L. Shue, G. Chen, M. Iseli, B. Gerratt, J. Neubauer, and A. Alwan, "Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation," *J. Acoust. Soc. Am.*, vol. 132, pp. 2625–2632, 2012.

[9] H. Fujisaki and M. Ljungqvist, "Proposal and evaluation of models for the glottal source waveform," in *ICASSP*, 1986, pp. 1605–1608.

[10] A. Rosenberg, "Effects of the glottal pulse shape on the quality of natural vowels," *J. Acoust. Soc. Am.*, vol. 49, pp. 583–590, 1971.

[11] G. Fant, *Acoustic theory of speech production*, 2nd ed. The Hague, Paris: Mouton, 1970, pp. 15-20.

[12] N. Henrich, C. d'Alessandro, and B. Doval, "Spectral correlates of voice open quotient and glottal flow asymmetry: theory, limits and experimental data," in *Eurospeech*, 2001, pp. 47–50.

[13] B. Story and I. Titze, "Voice simulation with a body cover model of the vocal folds," *J. Acoust. Soc. Am.*, vol. 97, pp. 1249–1260, 1995.

[14] I. Titze, "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.*, vol. 123, pp. 2733–2749, 2008.

[15] I. Titze, T. Riede, and P. Popolo, "Nonlinear source-filter coupling in phonation: Vocal exercises," *J. Acoust. Soc. Am.*, vol. 123, pp. 1902–1915, 2008.

[16] M. Zañartu, L. Mongeau, and G. R. Wodicka, "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.*, vol. 121, pp. 1119–1129, 2007.

[17] R. A. Samlan and B. H. Story, "Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling," *J. Speech Lang. Hear. Res.*, vol. 54, no. 5, pp. 1267–1283, 2011.

[18] I. R. Titze, *The myoelastic aerodynamic theory of phonation*. National Center for Voice and Speech, Denver, CO, 2006.

[19] I. Titze and B. Story, "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am.*, vol. 101, pp. 2234–2243, 1997.

[20] B. H. Story, "A parametric model of the vocal tract area function for vowel and consonant simulation," *J. Acoust. Soc. Am.*, vol. 117, pp. 3231–3254, 2005.

[21] G. Chen, J. Kreiman, Y.-L. Shue, and A. Alwan, "Acoustic correlates of glottal gaps," in *Interspeech*, 2011, pp. 2673–2676.

[22] G. Chen, Y.-L. Shue, J. Kreiman, and A. Alwan, "Estimating the voice source in noise," in *Interspeech*, 2012, pp. 1600–1603.

[23] M. R. Draper, B. Blagnys, and D. J. Premachandra, "To 'EE' or not to 'EE'," *J. Otolaryngology*, vol. 36, pp. 191–195, 2007.

[24] "GlotAnTools," [Computer software], Version 5, Department of Phoniatrics and Pediatric Audiology, University Hospital, Erlangen, Germany. Retrieved Nov. 11, 2011.

[25] G. Chen, J. Kreiman, B. R. Gerratt, J. Neubauer, Y.-L. Shue, and A. Alwan, "Development of a glottal area index that integrates glottal gap size and open quotient," *J. Acoust. Soc. Am.*, vol. 133, pp. 1656–1666, 2013.

[26] Y.-L. Shue, "VoiceSauce: a program for voice analysis," 2010, http://www.ee.ucla.edu/ spapl/voicesauce/ (Last viewed Apr. 1, 2012).

[27] H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigné, "Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based f0 extraction: Possible role of a repetitive structure in sounds," *Speech Commun.*, vol. 27, pp. 187–207, 1999.

[28] K. Sjölander, "The snack sound toolkit," KTH Stockholm, Sweden, 2004, http://www.speech.kth.se/snack/ (Last viewed Aug. 1, 2009).

[29] Y.-L. Shue and A. Alwan, "A new voice source model based on high-speed imaging and its application to voice source estimation," in *ICASSP*, 2010, pp. 5134–5137.

[30] I. R. Titze, "Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation," *J. Speech Lang. Hear. Res.*, vol. 49, pp. 439–447, 2006.

[31] G. Fant, "Preliminaries to analysis of the human voice source," *STL-QPSR*, vol. 4, pp. 1–27, 1982.

[32] B. Story, "LeTalker: Lumped-element model of phonation," 2012, http://sal.shs.arizona.edu/ bstory/LeTalkerMain.html (Last viewed Mar. 1, 2013).

[33] M. Hirano and D. M. Bless, *Videostroboscopic examination of the larynx*. Singular, San Diego, CA, 1993.

[34] S. Schuberth, U. Hoppe, M. Döllinger, J. Lohscheller, and U. Eysholdt, "High-precision measurement of the vocal fold length and vibratory amplitudes," *The Laryngoscope*, vol. 112, pp. 1043–1049, 2009.

[35] H. Imagawa, K.-I. Sakakibara, I. T. Tokuda, M. Otsuka, and N. Tayama, "Estimation of glottal area function using stereo-endoscopic high-speed digital imaging," in *Interspeech*, 2010, pp. 1005–1008.

[36] D. D. Mehta, M. Zañartu, T. F. Quatieri, D. D. Deliyski, and R. E. Hillman, "Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy," *J. Acoust. Soc. Am.*, vol. 130, pp. 3999–4009, 2011.

[37] Y.-L. Shue, G. Chen, and A. Alwan, "On the interdependencies between voice quality, glottal gaps, and voice-source related acoustic measures," in *Interspeech*, 2010, pp. 34–37.