



Published in final edited form as:

Ann Otol Rhinol Laryngol. 2010 January ; 119(1): 1–9.

Voice production mechanisms following phonosurgical treatment of early glottic cancer

Daryush D. Mehta, SM^{1,3,5}, Dimitar D. Deliyski, PhD⁴, Steven M. Zeitels, MD^{1,2}, Thomas F. Quatieri, ScD^{3,5}, and Robert E. Hillman, PhD^{1,2,3}

¹ Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital

² Department of Surgery, Harvard Medical School

³ Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology

⁴ Department of Communication Sciences and Disorders, University of South Carolina

⁵ Lincoln Laboratory, Massachusetts Institute of Technology

Abstract

Objectives—Although near-normal conversational voices can be achieved in the phonosurgical management of early glottic cancer, there are still acoustic and aerodynamic deficits in vocal function that must be better understood to help further optimize phonosurgical interventions. Stroboscopic assessment is inadequate for this purpose.

Methods—A newly-developed color ultra high-speed videoendoscopy (HSV) system that included time-synchronized recordings of the acoustic signal was used to perform a detailed examination of voice production mechanisms in 14 subjects. Digital image processing techniques were used to quantify glottal phonatory function and to delineate relationships between vocal fold vibratory properties and acoustic perturbation measures.

Results—Results for multiple measurements of vibratory asymmetry showed that 31–62 % of subjects displayed higher than normal average values, while mean values for glottal closure duration (open quotient) and periodicity of vibration fell within normal limits. Average HSV-based measures did not correlate significantly with acoustic perturbation measures, but moderate correlations were exhibited between the acoustic measures and standard deviations of the HSV-based parameters.

Conclusions—The use of simultaneous, time-synchronized HSV and acoustic recordings can provide new insights into post-operative voice production mechanisms that can not be obtained with stroboscopic assessment.

Keywords

voice production; glottic cancer; high-speed videoendoscopy; vocal fold asymmetry; acoustic perturbation

Correspondence: Robert E. Hillman, PhD, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, 1 Bowdoin Square, 11th Floor, Boston, MA 02114. Phone: 617-726-0220. Fax: 617-726-0222., hillman.robert@mgh.harvard.edu.
Reprint requests to be sent to Robert E. Hillman, PhD, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, 1 Bowdoin Square, 11th Floor, Boston, MA 02114.

Presented at the meeting of the American Broncho-Esophagological Association, Phoenix, Arizona, May 28–29, 2009.

Introduction

Early glottic cancer has been associated with high cure rates since direct endolaryngeal treatment techniques were established almost a century ago.¹ Since it is relatively easy to cure these neoplasms, voice outcome has always been a key metric of success of the treatment. In the mid 20th century, this concern about voice outcome led many surgeons to resect the thyroid lamina lateral to the tumor when a laryngofissure and cordectomy would have provided adequate oncologic treatment. The phonatory function of most cordectomy patients reflected aerodynamic glottal incompetence. By removing the thyroid cartilage,² the reconstructed soft tissue collapsed to form a competent neoglottic valve. However, at present, open laryngeal treatment approaches are seldom done as initial primary treatment.

As radiotherapy became the mainstay of treatment for early glottic cancer, those who espoused surgical management relied primarily on microlaryngeal treatment strategies. Most surgeons, however, lack confidence about consistent voice outcomes from endoscopic treatment, so radiotherapy remains as the dominant treatment strategy.³ To address this, we enhanced the voice outcome in treating early glottic cancer by performing narrow-margin resections for the more superficial lesions^{4,5} and by reconstructing the paraglottic space using the laryngeal framework using reconstructive methods⁶⁻⁸ transferred from skills acquired in managing paralytic dysphonia. These later techniques became the key strategy to reestablish phonatory aerodynamic competency. Most recently, we have been able to further diminish loss of soft tissue from the paraglottic region by employing 532 nm KTP angiolytic laser treatments.³ Although we have previously reported that near-normal conversational voices can be achieved in the phonosurgical management of early glottic cancer, there are still acoustic and aerodynamic deficits in vocal function that must be better understood to help further optimize phonosurgical interventions.^{3,9}

The purpose of this study was to gain better insights into the voice production mechanisms of patients who have undergone endoscopic phonosurgical treatment of early glottic cancer. The investigation was accomplished using a recently-developed system that could acquire digital color images of vocal fold vibration using high-speed videoendoscopy (HSV) with time-synchronized recordings of the acoustic voice signal during phonation. This enabled direct correlations between acoustic parameters and measures of glottal closure and vibratory symmetry extracted from the high-speed digital recordings using custom-designed image processing algorithms. The focus of the study was to determine the impact of vocal fold vibratory asymmetries and glottal closure on acoustic perturbation measures used in the clinic. Normative ranges for the HSV-based measures and acoustic parameters were derived from previous studies and provide a descriptive context for the current subject sample.

Methods

This study used HSV and acoustic voice data from fourteen subjects (12 male, 2 female) who had undergone endoscopic phonosurgical management of early glottic cancer. This was comprised primarily of angiolytic laser treatment with ultra-narrow margins which allowed for maximal preservation of paraglottic soft tissue. The average age of the 14 subjects in the current study was 66 years, with a range from 50 years to 84 years. Ten of the subjects were treated with the KTP angiolytic laser, two were treated with the PDL angiolytic laser, and two were treated with both types of lasers. HSV assessment was performed within a year of the most recent procedure for all the subjects, except two, who were treated over two years prior to data acquisition. Subjects were evaluated in a sound-treated room using transoral rigid endoscopy and instructed to sustain the vowel/i/ at a comfortable and steady pitch and loudness for two to four seconds.

HSV data acquisition

HSV data were acquired using the Phantom v7.3 high-speed video camera (Vision Research, Inc., Wayne, NJ), enabling high-quality color image capture at fast rates due to a sensitive CMOS image sensor. A C-mount lens adapter with adjustable focal length (KayPENTAX, Lincoln Park, NJ) was placed between the image sensor and a 70° transoral rigid endoscope (JEDMED, St. Louis, MO). Images were saved to partitions in a 4 GB on-board memory buffer and downloaded to a personal computer's hard drive via Ethernet after recordings were complete. HSV data were recorded at 4000 or 6250 images per second with maximum integration time and a spatial resolution of 320 horizontal × 480 vertical pixels to capture an approximately 2 cm² target area. The light source contained a short-arc Xenon lamp rated at 300 watts (KayPENTAX, Lincoln Park, NJ). The fan-cooled housing produced a collimated beam of light with a color temperature of over 6000 K. Three glass infrared (two dichroic, one absorbing) filters blocked infrared light to reduce thermal energy buildup during endoscopy.

Acoustic voice signal acquisition

The acoustic signal was recorded using a head-mounted, high-quality condenser microphone (Model MKE104, Sennheiser electronic GmbH, Wennebostel, Germany) with a cardioid pattern, offering directional sensitivity and a wideband frequency response. The microphone was situated approximately 4 cm from the lips at a 45° azimuthal offset. The microphone's gain circuitry (Model 302 Dual Microphone Preamplifier, Symetrix, Inc., Mountlake Terrace, WA) offered a low-noise, low-distortion input signal to additional preconditioning electronics (CyberAmp Model 380, Axon Instruments, Inc., Union City, CA) for gain control and anti-aliasing using low-pass filtering at a 3 dB cutoff frequency of 30 kHz. The analog signals were digitized at a 100 kHz sampling rate, 16-bit quantization, and a ±10 V dynamic range by a PCI digital acquisition board (6259 M series, National Instruments, Austin, TX).

Time-synchronization of HSV data and microphone signal

Time-synchronization of the HSV data and the microphone signal was critical for enabling correlations between voice production measures and acoustic perturbation from the same phonatory segment. The camera clock was supplied by the National Instruments board's clock source that was synchronized to the sampling of the acoustic signal. The hardware clock division and data acquisition settings were controlled by MiDAS DA software (Xcitex Corporation, Cambridge, MA). Alignment of the HSV data and the microphone signals was accomplished by recording an analog signal from the camera that precisely indicated the time of the last recorded image. To compensate for the larynx-to-microphone acoustic propagation time, the microphone signal was shifted into the past by 600 μs relative to the HSV data. This shift duration was approximated from a propagation length of 21 cm (17 cm vocal tract + 4 cm lip-to-microphone distance) and a speed of sound of 35 400 cm/s. Precise compensation for differences in the vocal tract lengths of each subject was not necessary for the goals of this study since data were based on averages derived from period-to-period measures across entire phonatory segments. Future investigations into acoustic-physiological correlations on a cycle-by-cycle basis would benefit from a more rigorous compensation scheme.

Signal processing of HSV data

Vocal fold tissue motion during phonation was measured by tracking the mediolateral motion of the left and right vocal fold edges closest to the glottal midline. The tracked edges collapsed the three-dimensional motion of the vocal folds that constituted the mucosal wave to a two-dimensional space that ignored inferior-superior phase differences.

In this study, the acoustic signal and the high-speed video images were segmented into approximately 400 ms segments for each subject exhibiting stable phonatory characteristics.

Semiautomatic algorithms improved upon manual markings of digital kymograms¹⁰ and included knowledge-based image processing of the kymograms and high-speed video images to extract vibratory asymmetry measures that were compared with properties of the acoustic voice signal. For each segment, vocal fold amplitude waveforms were obtained from the images using the algorithm illustrated in Figure 1.

HSV-based measures depended on accurate extraction of vocal fold tissue motion extracted by demarcating the edges of the digital kymograms. Motion artifacts were compensated for^{11,12} so that edge detection was equivalent to motion tracking of the medial vocal fold edge. The anterior-posterior axis (glottal midline) was defined interactively on the first image capturing maximum vocal fold abduction. The midline further defined the line of symmetry connecting the anterior commissure to the posterior end of the musculo-membranous vocal folds. All HSV frames were rotated such that the glottal midline was oriented vertically, and a cropping stage framed the glottis to reduce processing time and constrain the image to the most relevant features for accurate edge detection. Using threshold-based edge detection, vocal fold amplitude traces were derived from digital kymograms taken at the midpoint of the membranous glottis, i.e., midway between the anterior commissure and the posterior end of the membranous glottis.

Objective measures of vocal fold vibratory asymmetry. HSV-based measures included left-right phase asymmetry, left-right amplitude asymmetry, frequency differences, and axis shifts during closure. vec and colleagues have validated this categorization by observing these different types of vibratory asymmetries in videokymographic images recorded from subjects with various vocal pathologies.¹³ Quantitative measures were derived to build on these findings and other studies that computed indices of left-right asymmetries from kymographic data.^{10, 14} Analyses were restricted to phase differences, amplitude differences, and axis shifts during closure because of the overall interest in revealing co-variations of physiological properties with synchronous acoustic data.

HSV-based measures

Five HSV-based measures were automatically extracted from the digital kymographs. **Left-right amplitude asymmetry** quantified the difference in the maximum excursions of the left and right vocal fold amplitude traces within a period. The zero amplitude reference was defined as the vocal fold position at the onset of its opening phase. **Left-right phase asymmetry** quantified temporal delays between maximum excursion of the vocal folds. The time delay was normalized by the period duration to obtain a percentage.¹⁰ The degree of **axis shift during closure** calculated the spatial distance, in the mediolateral axis, between onsets of closure and subsequent opening phases, and yielded a percentage after normalization by the glottal width. The HSV-based **open quotient** was the ratio, in percent, between the open phase duration and the entire period. For each of the above cycle-to-cycle measures, mean and standard deviation values were obtained over the number of cycles in the phonatory segment. Finally, **period irregularity** was computed from the digital kymogram to reflect jitter-like phenomenon. Period durations were defined as the difference between successive onsets of glottal closure. The absolute difference between successive period durations was calculated, and the period irregularity percentage yielded the average of these absolute differences divided by the mean period duration.

Acoustic perturbation measures

Standard acoustic measures of **shimmer**, **jitter**, and **noise-to-harmonics** ratio were computed from the synchronously-recorded acoustic voice signal using the Multi-Dimensional Voice Program (MDVP, Model 5105, KayPENTAX, Lincoln Park, NJ). Percent shimmer was defined as the average absolute difference between peak amplitudes of consecutive periods

divided by the average peak amplitudes during the phonatory segment. Similarly, percent jitter was defined as the average absolute difference between successive periods divided by the average period duration. The noise-to-harmonics ratio (NHR) reflected the relative spectral energy contributions of the noise component and the harmonic component of the acoustic voice signal.

Comparison with normative data

Limited data sets exist to make extensive comparisons among subjects with and without vocal pathology. Normative values for HSV-based measures have only been published for the left-right phase asymmetry parameter based on manual measurements of digital kymograms from 52 subjects.¹⁰ The average age of these subjects was 41 years, with a range from 18 years to 65 years. The subjects were age- and sex-matched, with 24 male and 28 female subjects divided among three age ranges: 18–33 years, 34–49 years, and 50–65 years. Normative values for the other HSV-based measures were obtained by re-analyzing the HSV recordings from that normative investigation with the signal processing methods described above. Averaged values (means, standard deviations, ranges) for the entire normal cohort were used to provide a descriptive context for the data obtained from the post-surgical subjects because statistical testing showed no age- or sex-based differences in measures within the normal group.

The normative range for open quotient was defined as the mean (μ) \pm 2 standard deviations (σ). For all other HSV-based measures, the normative range was defined as values less than or equal to 2 standard deviations above the mean. Normal ranges of measures derived from the normative study were included to provide a descriptive context for which HSV-based measures of the subjects in the current study could be compared. A more rigorous analysis inferring statistical differences between subjects with and without vocal pathology is a subject for future work in this area. Normative ranges for the acoustic perturbation measures were taken from standard clinical assessment thresholds reported by MDVP.

Results

Data from one (male, 64 years old) of the 14 subjects was eliminated prior to further statistical assessment of group results because the acoustic signal that was recorded during endoscopy was too unstable for standard perturbation analyses (exhibiting Type II signal characteristics: sudden qualitative changes in the form of subharmonics or modulation frequencies that interfere with the fundamental frequency¹⁵). Auditory-perceptual ratings were documented for a majority of the subjects by a certified speech-language pathologist on staff using the CAPE-V protocol.¹⁶ Voices fell within the mild-to-moderate range for overall severity of dysphonia, agreeing with our previous observations that, although near-normal conversational voices can be achieved, persistent perceptual acoustic deficits require additional investigation.

Table 1 shows a summary of HSV and acoustic measures for 13 patients relative to normative values. As described above, normative values for the HSV data were obtained by analyzing recordings from a previous investigation of 52 normal subjects.¹⁰ Means, standard deviations, and normative ranges for acoustic perturbation measures were taken from standard clinical assessment thresholds reported by MDVP.

Distributions of all the HSV-based measures for the patient group are displayed in Figure 2, with indications of the normal ranges that were defined above. Five out of the 13 subjects (38 %) exhibited left-right phase asymmetry values that were above the normative threshold. Eight out of the 13 subjects (62 %) exhibited abnormal amplitude asymmetry values, while 4 out of the 13 subjects (31 %) exhibited axis shift values above the normative threshold. None of the subjects displayed average open quotient or period irregularity values that were outside of the normative range. The average values for the acoustic measures (see Table 1) were comparable

to those from previous reports on vocal function in this patient group,^{3,9} with a post-operative persistence of elevations in jitter and shimmer and, essentially, normal NHR values.

Due to the non-normal marginal distributions of all the measures and the limited number of subjects in the sample (violating the assumptions for linear correlation analysis using Pearson's r), nonparametric statistical analysis using the Spearman rank-order correlation method was performed. Spearman ρ was computed from transformed, rank-ordered measures. Pairwise correlation analysis was performed between HSV-based measures and acoustic voice parameters. Although none of the correlations between average values of the measures was significant, statistically significant correlations ($p \leq .05$) were found between the standard deviations of the HSV-based measures and acoustic perturbation parameters in three instances (see Table 2). Scatter plots with corresponding regression lines for each of these statistically-significant correlations are depicted in Figure 3.

As Table 2 reports, jitter in the acoustic voice signal significantly correlated with the HSV-based measure reflecting the standard deviation of left-right phase asymmetry. Spearman ρ for this correlation was 0.604, indicating that approximately 36 % of the variance (i.e., ρ^2) in jitter was explained by the variation in left-right phase asymmetry. Acoustic jitter also correlated significantly with the standard deviation of left-right amplitude asymmetry, with a correlation coefficient ρ of 0.621. Thus, approximately 39 % of the variance in jitter was accounted for by variation in the left-right amplitude asymmetry measure. Finally, shimmer in the acoustic voice signal was most significantly correlated ($\rho = 0.659$) with the HSV-based standard deviation of the open quotient, corresponding to an explained variance of 43 %. The noise-to-harmonics ratio did not correlate significantly with any of the HSV-based measures computed.

Discussion

This study used a newly-developed HSV-based system to examine voice production mechanisms in patients who had undergone phonosurgical management of glottic cancer. The analyses focused on quantitative measures of glottal closure and vibratory symmetry (i.e., vocal fold kinematics) extracted via custom-designed digital image processing routines and the correlations of these measures with acoustic perturbation parameters.

Taken together, the HSV and acoustic measurement results for this patient group described postoperative voice production mechanisms in which glottal closure (open quotient) and the overall periodicity of tissue vibration during phonation were within the normal ranges that were defined for this study. The restoration and/or maintenance of normal-like glottal closure durations and overall vibratory periodicity probably contributed to the normal acoustic NHR values that have also been observed previously in similar groups of patients.³⁻⁹ On the other hand, the voice production mechanisms of these patients also displayed higher than normal within-cycle asymmetries in vocal fold tissue vibration (phase, amplitude, and closure axis) and, as previously reported,³⁻⁹ post-operative persistence of elevations in acoustic jitter and shimmer.

The results of pairwise correlations between average values for HSV and acoustic perturbation measures revealed that variations in levels of acoustic jitter and shimmer were not related to within-cycle asymmetries in tissue vibration. Instead, a significant amount of the variation in acoustic jitter was accounted for by the changes (standard deviations) in the symmetry of the phase and amplitude of tissue motion across vibratory cycles, and a significant amount of the shimmer variation was correlated with changes (standard deviation) in glottal closure durations (open quotients) across cycles of vibration. Thus, it appeared that even though open quotient values for patients fell within the normal range that was defined for this study ($\text{mean} \pm 2$

standard deviations), this range was quite broad, with the larger variations within this range actually being associated with elevated levels of acoustic shimmer.

Time-varying asymmetries were expected to play a large role in affecting characteristics of the acoustic voice signal. In a recent study, the physiological correlates of acoustic perturbation measures were investigated by simulating asymmetric vocal fold vibrations using a mathematical model.¹⁷ It was shown that time-varying changes in left-right asymmetry resulted in acoustic jitter and shimmer, whereas time-invariant left-right asymmetries yielded an acoustic output that did not exhibit increased values of jitter and shimmer. This conclusion was supported by the results of the current study.

The apparently critical cycle-to-cycle variations in tissue vibratory behavior that were shown to be significantly correlated with the degradation of the acoustic signal in this investigation would not be reliably revealed with stroboscopy. In fact, since stroboscopy would only provide stable imaging of periodic (repeating) vocal fold motion, it would be capable of only capturing the kind of highly-repetitive asymmetries that do not appear to make a major contribution to disrupting acoustic sound generation, assuming that glottal closure is adequate.

The results of this study indicated that the post-operative vocal function of this patient group could be further improved by reducing cycle-to-cycle variations in the phase relationships of vibratory tissue. This might be achievable if it were possible to restore lost pliability to specific regions of the vocal folds, depending on the nature and extent of the resection that was needed to ameliorate the cancer. The need to restore pliability in such cases is one of the major motivations behind recent efforts to develop vocal fold phonatory mucosal implants.¹⁸

The post-resection healed glottal epithelium assumes the biomechanical properties of the underlying scar and/residual soft tissue of the neocord.⁹ Therefore, the overall goal of a superficial lamina propria (SLP) replacement implant would be to more closely equilibrate the biomechanical properties within and between the vocal folds and thereby reduce sources of cycle-to-cycle asymmetric vibration. Given the limitations described above for stroboscopy, it is clear that HSV imaging will provide insights that will assist surgeons in geographically mapping lost phonatory mucosal pliability to target restoration as well as assess post-reconstruction tissue rheology and its impact on enhancing vocal function.

The HSV data in this study was limited to the mid-portion of the musculo-membranous glottis. This is the typical location for performing videokymography, which is only capable of recording activity at one location.¹³ The differential variation in tissue pliability that can occur within and between vocal folds following surgical management of glottic cancer can give rise to asymmetries in tissue motion at other points along the glottis. The digital implementation of kymography following post-HSV image capture that was used in this study has the advantage of being able to generate kymographic images (slices) at any location along the entire length of the glottis, thus providing the potential to isolate motion at specific locations of interest. This capability is demonstrated in Figure 4 where digital kymograms were extracted from three glottal locations. Efforts are currently underway to determine how such multiple measures can be combined/presented in ways that are meaningful and allow useful comparisons with other simultaneously-gathered (time-synchronized) measures of vocal function, including acoustic, electroglottographic, neck skin acceleration, and aerodynamic parameters.

Even though significant correlations were found between the standard deviations of some HSV measures and acoustic perturbation parameters, the actual correlations were at a moderate level. This means that there is still a sizable amount of variation in acoustic perturbation that is not accounted for by the HSV measures used in this study. It is hoped that more complete analysis of the HSV imaging data, such as the use of multiple kymograms described above, will increase the amount of acoustic perturbation that can be accounted for. However, it is also possible that

there are other sources of acoustic perturbation in addition to asymmetries in vocal fold tissue motion, including higher-order aerodynamic phenomenon/interactions,^{19,20} that may play a role during voice production.

Conclusion

The use of simultaneous, time-synchronized HSV and acoustic recordings can provide new insights into the post-operative voice production mechanisms of patient who have undergone phonosurgical management of vocal pathology, including glottic cancer. In particular, this approach allows for the application of automated digital image and signal processing methods to reveal critical relationships between vocal fold vibratory function and the resulting acoustic characteristics of the voice. Such investigations are not possible with standard clinical stroboscopy methods. This information should substantially assist surgeons in identifying biomechanical phonatory mucosal deficits and the effectiveness of implant reconstruction efforts.

Even though HSV provides much more detailed information about vocal fold phonatory function than stroboscopy, its eventual adoption into standard clinical practice will depend on the extent to which remaining practical, technical, and methodological, challenges can be met. Such challenges include the relatively high cost of HSV systems, management of large computer files that HSV recordings produce, limitations on the sampling of vocal behaviors due to the brief durations of HSV recordings, and a paucity of solid clinical research that demonstrates HSV significantly improves the diagnosis and management of voice disorders (i.e., clinical trials with controls). Work is already well underway to address many of the remaining barriers to the clinical adoption of HSV.¹²

Acknowledgments

Supported by grants from the NIH National Institute on Deafness and Other Communication Disorders (T32 DC00038 and R01 DC007640) and by the Institute for Laryngology and Voice Restoration. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The work of T. F. Quatieri was sponsored under Air Force Contract 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.

References

1. Lynch RC. Intrinsic carcinoma of the larynx, with a second report of the cases operated on by suspension and dissection. *Trans Am Laryngol Assoc* 1920;40:119–26.
2. Som ML. Hemilaryngectomy a modified technique for cordal carcinoma with extension posteriorly. *Arch Otolaryngol* 1951;54:524–33.
3. Zeitels SM, Burns JA, Lopez-Guerra G, Anderson RR, Hillman RE. Photoangiolytic laser treatment of early glottic cancer: a new management strategy. *Ann Otol Rhinol Laryngol Suppl* 2008;117(suppl 199):1–24.
4. Zeitels SM. Premalignant epithelium and microinvasive cancer of the vocal fold: the evolution of phonomicrosurgical management. *Laryngoscope* 1995;105:1–51. [PubMed: 7885166]
5. Zeitels SM. Phonomicrosurgical treatment of early glottic cancer and carcinoma in situ. *Am J Surg* 1996;172:704–9. [PubMed: 8988684]
6. Hartig G, Zeitels SM. Optimizing voice in conservation surgery for glottic cancer. *Oper Techn Otolaryngol Head Neck Surg* 1998;9:214–23.
7. Zeitels SM, Jarboe J, Franco RA. Phonosurgical reconstruction of early glottic cancer. *Laryngoscope* 2001;111:1862–5. [PubMed: 11801959]
8. Zeitels SM. Optimizing voice after endoscopic partial laryngectomy. *Otolaryngol Clin North Am* 2004;37:627–36. [PubMed: 15163606]

9. Zeitels SM, Hillman RE, Franco RA, Bunting GW. Voice and treatment outcome from phonosurgical management of early glottic cancer. *Ann Otol Rhinol Laryngol Suppl* 2002;111(suppl 190):1–20.
10. Bonilha HS, Deliyski DD, Gerlach TT. Phase asymmetries in normophonic speakers: visual judgments and objective findings. *Am J Speech Lang Pathol* 2008;17:367–76. [PubMed: 18840697]
11. Deliyski DD. Endoscope motion compensation for laryngeal high-speed videoendoscopy. *J Voice* 2005;19:485–96. [PubMed: 16102674]
12. Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. Clinical implementation of laryngeal high-speed videoendoscopy: challenges and evolution. *Folia Phoniatri Logop* 2008;60:33–44. [PubMed: 18057909]
13. Svec JG, Sram F, Schutte HK. Videokymography in voice disorders: what to look for? *Ann Otol Rhinol Laryngol* 2007;116:172–80. [PubMed: 17419520]
14. Qiu Q, Schutte HK, Gu L, Yu Q. An automatic method to quantify the vibration properties of human vocal folds via videokymography. *Folia Phoniatri Logop* 2003;55:128–36. [PubMed: 12771464]
15. Titze, IR. Workshop on acoustic voice analysis: summary statement; Denver, CO: National Center for Voice and Speech; 1995. p. 1-36.
16. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* 2009;18:124–32. [PubMed: 18930908]
17. Nardone, M. Master of Science thesis, Physics. Bowling Green State University; 2007. Analysis of voice perturbations using an asymmetric model of the vocal folds.
18. Zeitels SM, Blitzler A, Hillman RE, Anderson RR. Foresight in laryngology and laryngeal surgery: a 2020 vision. *Ann Otol Rhinol Laryngol Suppl* 2007;116(suppl 198):1–16.
19. Murugappan S, Khosla S, Casper K, Oren L, Gutmark E. Flow fields and acoustics in a unilateral scarred vocal fold model. *Ann Otol Rhinol Laryngol* 2009;118:44–50. [PubMed: 19244963]
20. McGowan RS, Howe MS. Compact green's functions extend the acoustic theory of speech production. *J Phon* 2006;35:259–70.

Over 4,000 high-speed video images are captured per second.

Three of the images are displayed here to illustrate the following analysis steps.

From one image with maximal glottal area, the user defines the **glottal midline** by indicating the anterior commissure and posterior vibrating end.

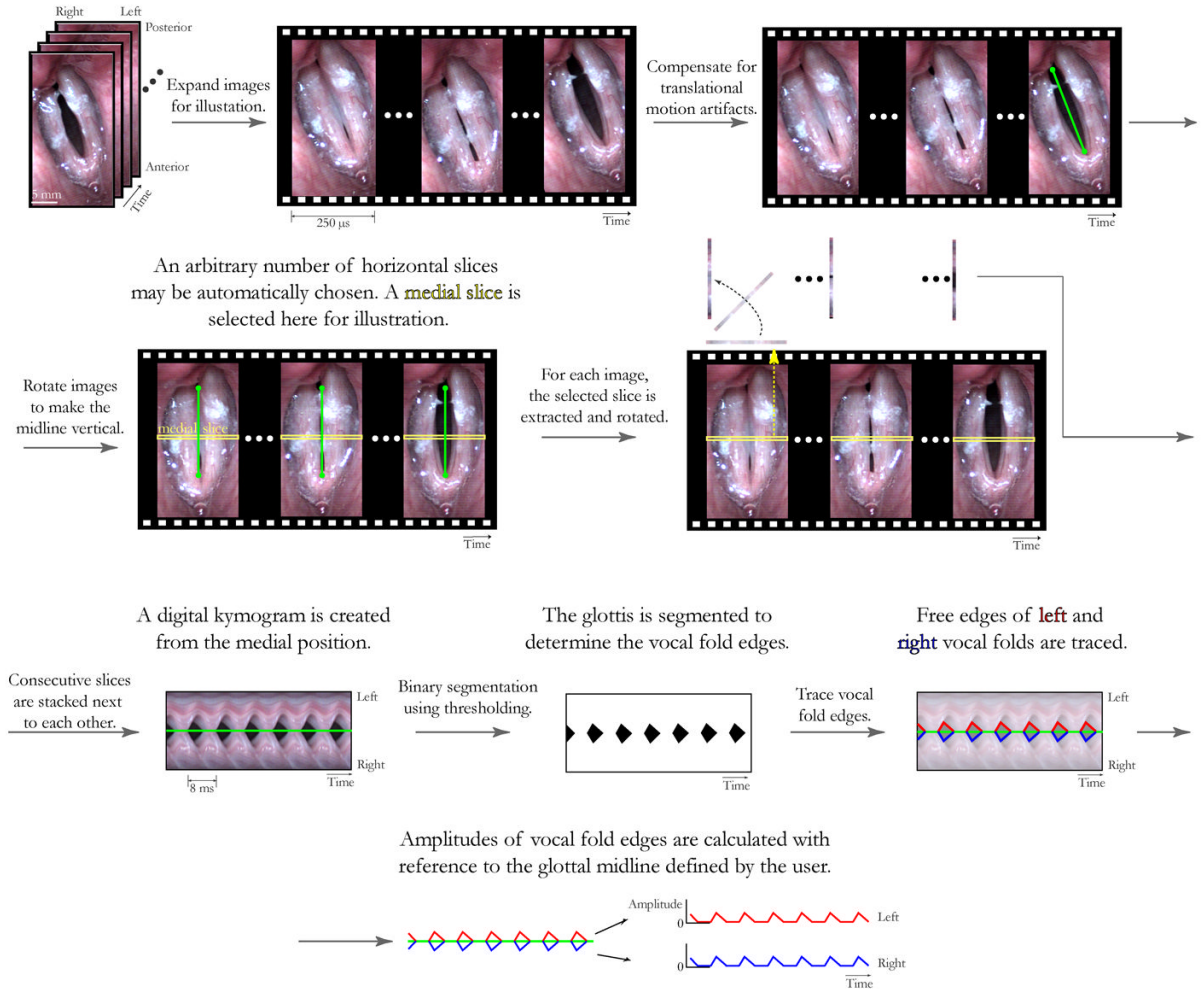


Figure 1. Algorithm for generating vocal fold amplitude waveforms from which all HSV-based measures are derived.

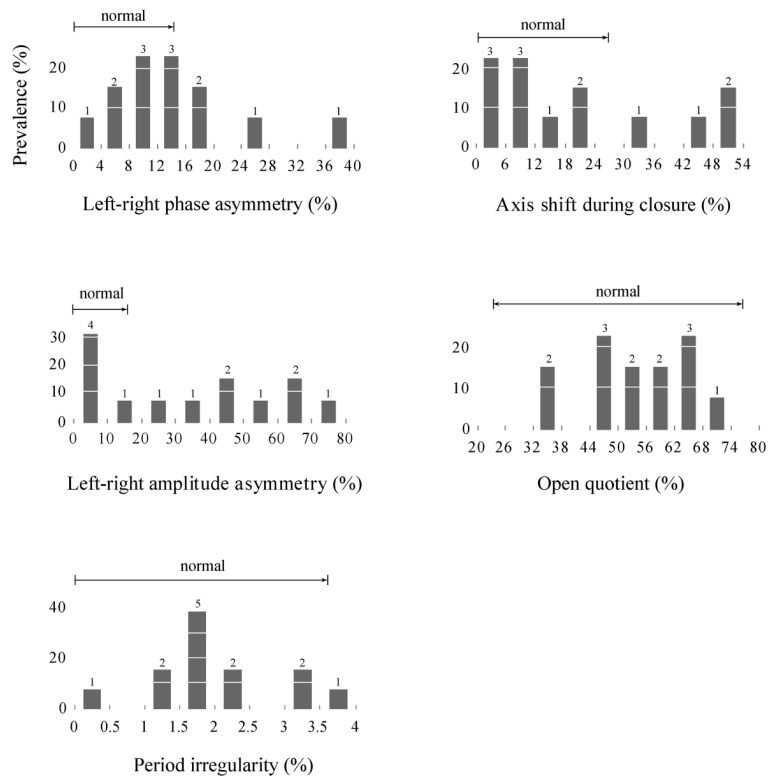


Figure 2. Distributions of the HSV-based measures for the patient group (N = 13) with normal ranges indicated (see text for definition of normal ranges).

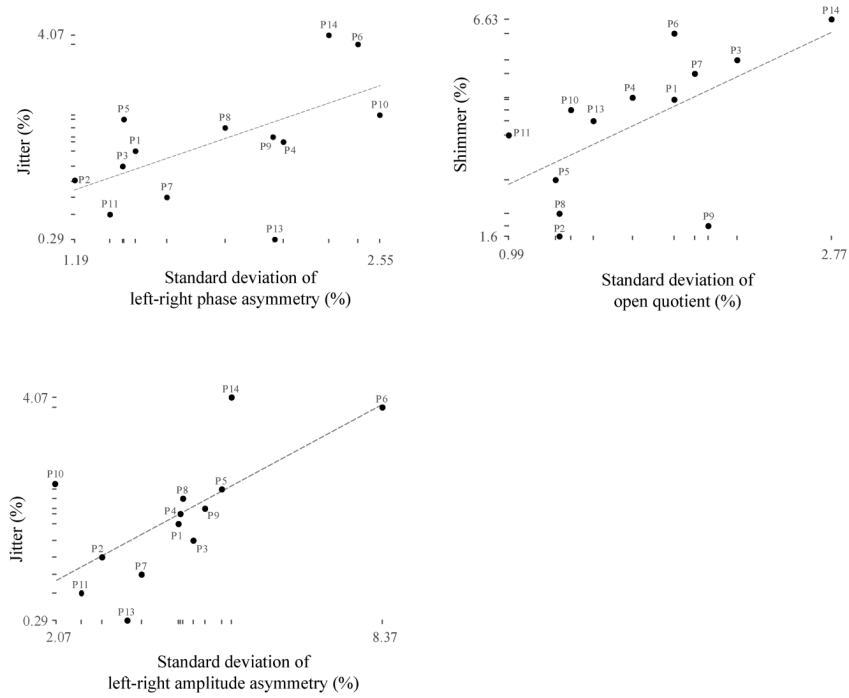


Figure 3. Scatter plots displaying relationships between HSV-based and acoustic measures exhibiting significant pairwise correlations.

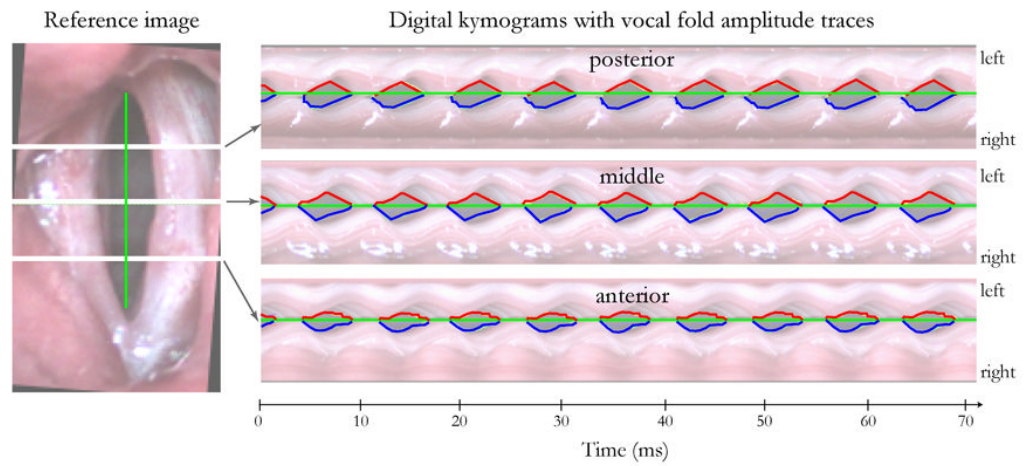


Figure 4. Illustration of deriving vocal fold amplitude waveforms from digital kymograms extracted from three locations along the anterior-posterior glottal axis.

Summary of HSV and acoustic measures for the patient group (N = 13). Normative values for HSV measures were obtained by analyzing recordings from 52 normal subjects.¹⁰ Means, standard deviations, and normative values for acoustic measures are from MDVP.

Table 1

| Measure | Normative sample | | | Subject sample | | |
|--|------------------|----------|--------------|----------------|----------|--|
| | μ | σ | Normal range | μ | σ | |
| HSV-based left-right phase asymmetry (%) | 6.27 | 4.33 | ≤ 14.9 | 14.1 | 9.10 | |
| HSV-based left-right amplitude asymmetry (%) | 6.45 | 4.66 | ≤ 15.8 | 34.1 | 26.5 | |
| HSV-based axis shift during closure (%) | 10.4 | 8.30 | ≤ 27.0 | 21.4 | 17.4 | |
| HSV-based open quotient (%) | 49.8 | 13.4 | [23.0, 76.6] | 54.7 | 11.4 | |
| HSV-based period irregularity (%) | 1.84 | 0.880 | ≤ 3.60 | 1.91 | 0.882 | |
| Acoustic jitter (%) | 0.616 | 0.434 | ≤ 1.04 | 2.06 | 1.10 | |
| Acoustic shimmer (%) | 2.21 | 0.915 | ≤ 3.81 | 4.21 | 1.67 | |
| Acoustic noise-to-harmonics ratio | 0.116 | 0.0123 | ≤ 0.190 | 0.143 | 0.0246 | |

Table 2

Significant pairwise correlations between HSV-based measures and acoustic measures (N = 13).

| Pairwise comparison | Spearman ρ | p-value | 95 % confidence interval for ρ |
|--|-----------------|---------|-------------------------------------|
| Acoustic jitter – Standard deviation of left-right phase asymmetry | 0.604 | 0.03 | (0.080, 0.867) |
| Acoustic jitter – Standard deviation of left-right amplitude asymmetry | 0.621 | 0.02 | (0.106, 0.873) |
| Acoustic shimmer – Standard deviation of open quotient | 0.659 | 0.01 | (0.170, 0.888) |