JSLHR

Research Article

Glottal Aerodynamic Measures in Women With Phonotraumatic and Nonphonotraumatic Vocal Hyperfunction

Víctor M. Espinoza,[a,b] Matías Zañartu,[a] Jarrad H. Van Stan,[c,d,e] Daryush D. Mehta,[c,d,e] and Robert E. Hillman[c,d,e]

Purpose: The purpose of this study was to determine the validity of preliminary reports showing that glottal aerodynamic measures can identify pathophysiological phonatory mechanisms for phonotraumatic and nonphonotraumatic vocal hyperfunction, which are each distinctly different from normal vocal function.
Method: Glottal aerodynamic measures (estimates of subglottal air pressure, peak-to-peak airflow, maximum flow declination rate, and open quotient) were obtained noninvasively using a pneumotachograph mask with an intraoral pressure catheter in 16 women with organic vocal fold lesions, 16 women with muscle tension dysphonia, and 2 associated matched control groups with normal voices. Subjects produced /pae/ syllable strings from which glottal airflow was estimated using inverse filtering during /ae/ vowels, and subglottal pressure was estimated during /p/ closures. All measures were normalized for sound pressure level (SPL) and statistically tested for differences between patient and control groups.
Results: All SPL-normalized measures were significantly lower in the phonotraumatic group as compared with measures in its control group. For the nonphonotraumatic group, only SPL-normalized subglottal pressure and open quotient were significantly lower than measures in its control group.
Conclusions: Results of this study confirm previous hypotheses and preliminary results indicating that SPL-normalized estimates of glottal aerodynamic measures can be used to describe the different pathophysiological phonatory mechanisms associated with phonotraumatic and nonphonotraumatic vocal hyperfunction.

Vocal hyperfunction (VH) refers to "chronic conditions of abuse and/or misuse of the vocal mechanism due to excessive and/or 'imbalanced' muscular forces" (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989, p. 373) and is associated with the most frequently occurring types of voice disorders. Hillman et al. (1989) proposed that two manifestations of VH reflect different underlying pathophysiological mechanisms that were originally referred to as *adducted* VH and *nonadducted* VH and more recently relabeled *phonotraumatic* VH (PVH) and *nonphonotraumatic* VH (NPVH; Mehta et al., 2015). PVH is associated with the formation of benign vocal fold lesions due to chronic tissue trauma (e.g., vocal fold nodules). NPVH is associated with chronic dysphonia and vocal fatigue in the absence of vocal fold tissue trauma or other conditions that could affect phonation and is often referred to as *primary muscle tension dysphonia* (MTD; Bhattacharyya, 2014).

In the view of Hillman et al. (1989), both PVH and NPVH involve increased tension and stiffness of the vocal folds due to heightened and/or imbalanced (uncoordinated) levels of laryngeal muscle activity with an associated increase in aerodynamic forces required to produce phonation. The two conditions are hypothesized to each distinctly differ from normal vocal function primarily in terms of the impact of VH on vocal fold adduction and abduction. In PVH, adduction forces appear to predominate to maintain tight approximation of the vocal folds and, in combination with increased aerodynamic parameters, create higher vocal fold collision forces and tissue trauma. In NPVH, an apparent imbalance between adduction and abduction forces precludes tight approximation of the vocal folds, thus

[a]Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
[b]Department of Sound, Universidad de Chile, Santiago
[c]Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston
[d]MGH Institute of Health Professions, Boston, MA
[e]Harvard Medical School, Boston, MA
Correspondence to Matías Zañartu: matias.zanartu@usm.cl

Editor: Julie Liss
Associate Editor: Nelson Roy
Received August 24, 2016
Revision received February 11, 2017
Accepted March 12, 2017
https://doi.org/10.1044/2017_JSLHR-S-16-0337

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

actually reducing the potential for trauma to vocal fold tissue even though heightened aerodynamic forces are necessarily used to produce phonation.

The original 1989 publication that proposed the two types of VH also included results from an initial study involving relatively small numbers of heterogeneous patients and generic controls that provided preliminary support for the view that the pathophysiological mechanisms associated with these two forms of VH are each quite different from normal phonatory mechanisms (Hillman et al., 1989). Further support for the distinguishing characteristics of the PVH variant was provided in three subsequent studies in subjects with vocal fold lesions (Hillman et al., 1990; Holmberg, Doyle, Perkell, Hammarberg, & Hillman, 2003; Sapienza & Stathopoulos, 1994). These earlier investigations used a combination of noninvasive acoustic and aerodynamic measures that were designed to provide insight into underlying pathophysiological mechanisms. It was believed that such measures would have the potential to improve the clinical management of these disorders by providing the first quantitative metrics upon which to base judgments about the type and severity of VH and by offering a more objective basis for treatment decisions (e.g., decisions about when to end voice therapy on the basis of a measurable reduction in VH).

Several measures were extracted from recordings of the acoustic signal, intraoral air pressure, and high-bandwidth oral airflow (pneumotachograph mask) during the repeated production of /pae/ syllables. Primary measures included acoustic measures of fundamental frequency (f_o) and sound pressure level (SPL), estimates of subglottal air pressure (SGP) from the intraoral air pressure, and estimates of glottal airflow waveform parameters extracted from the oral airflow volume velocity (OVV) using inverse filtering (IF). Glottal airflow measures included peak-to-peak amplitude of the unsteady airflow (also denoted as AC flow [ACFL]), maximum flow declination rate (MFDR), and open quotient (OQ). In general, patients with PVH displayed abnormally elevated values for SGP, ACFL, and MFDR, which was interpreted to reflect increased potential for trauma to vocal fold tissue that contributed to the chronic presence of vocal fold lesions and associated dysphonia. Patients with PVH also tended to have elevated OQ values that were attributed to the obstruction of glottal closure due to the presence of vocal fold lesions (Hillman et al., 1989, 1990). Patients with NPVH also displayed abnormally elevated values for SGP and OQ without concomitant increases in ACFL and MFDR, which was associated with inefficient phonation and dysphonia but decreased potential to cause trauma to vocal fold tissue.

In addition, aerodynamic measures appeared to be more sensitive to the presence of vocal pathology than acoustic measures, particularly when the aerodynamic measures were normalized with respect to SPL. Preliminary results from the literature were supported by a more recent attempt to computationally model PVH that predicted increases in ACFL and MFDR and vocal fold collision forces when increasing SGP in the presence of incomplete glottal closure to maintain a given SPL (Zañartu et al., 2014). This result essentially confirms what was observed in the 1989 data from patients with PVH and is interpreted as reflecting the potential role of compensation (secondary or reactive VH) in perpetuating phonotrauma (i.e., a "vicious cycle"; Hillman et al., 1989, p. 383). Even though the early results reflecting different (quantifiable) pathophysiological mechanisms for PVH and NPVH were promising (and partially corroborated for PVH with modeling), these findings have never been completely validated in studies on homogeneous groups of patients with hyperfunctional voice disorders and well-matched controls that are large enough for formal statistical testing, nor have objective glottal airflow measures been adopted for use in routine clinical assessment or therapy.

The purpose of the present study was to determine whether preliminary evidence that glottal aerodynamic measures can identify distinct pathophysiological mechanisms for PVH and NPVH (Hillman et al., 1989) can be confirmed in homogeneous groups of patients and well-matched normal control subjects. The two primary hypotheses tested were that (a) patients with PVH (vocal fold nodules or polyps) would display significantly lower values for the SPL-normalized aerodynamic parameters of SGP, ACFL, MFDR, and OQ than would be found in matched controls with normal voices and (b) patients with NPVH (primary MTD) would display significantly lower values for SGP and OQ only than would be seen in matched controls with normal voices. Positive findings would support the capability of the measurement approach to objectively quantify the pathophysiological mechanisms associated with PVH and NPVH and justify further development and application of these measures for clinical use (e.g., as outcome measures and biofeedback targets during voice therapy).

Method

Participants

Two groups of adult female subjects with voice disorders were analyzed: 16 patients with PVH (vocal fold nodules or polyps) and 16 patients with NPVH (primary MTD). Diagnoses were based on a complete team evaluation by laryngologists and speech-language pathologists at the Massachusetts General Hospital Voice Center that included (a) a complete case history, (b) endoscopic imaging of the larynx, (c) aerodynamic and acoustic assessment of vocal function, (d) a patient-reported Voice-Related Quality of Life (V-RQOL) questionnaire (Hogikyan & Sethuraman, 1999), and (e) clinician-administered Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) assessment (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer, & Hillman, 2009). All patients were enrolled prior to the administration of any voice treatment. Table 1 summarizes demographics of the patients. Table 2 shows group-based averages for the V-RQOL and CAPE-V assessments.

Each patient aided in identifying her own control subject with normal vocal status who was matched for sex,

Table 1. Demographics of the two patient groups: phonotraumatic (nodules and polyps) and nonphonotraumatic (muscle tension dysphonia [MTD]) vocal hyperfunction.

Occupation	No. subject pairs	Patient diagnoses
Actor	2	MTD
Administrator	3	MTD (2), nodules (1)
Admin. assistant	1	MTD
College student	2	MTD
Consultant	2	Nodules (1), polyp (1)
Event planner	1	Polyp
Fitness instructor	2	MTD (1), nodules (1)
Full-time mother	3	MTD
Marketing	1	Nodules
Media relations	1	Nodules
Music teacher	1	Nodules
Psychologist	1	Nodules
Registered nurse	3	MTD (1), nodules (1), polyp (1)
Sales	1	Polyp
Social worker	1	MTD
Systems analyst	1	MTD
Talent recruiter	2	Nodules
Teacher	4	MTD (2), nodules (2)

occupation, and approximate age (± 5 years). The normal vocal status of all 32 control subjects was verified by a licensed speech-language pathologist specializing in voice disorders via interview (subjects reported no difficulties with their voices in daily life), laryngeal videostroboscopic examination, and CAPE-V assessment. The ages (mean ± standard deviation) of the PVH and matched control groups were 32.3 ± 12.8 years and 32.9 ± 12.9 years, respectively, and thus not statistically different. Similarly, the ages of the NPVH and matched control groups were not statistically different at 42.1 ± 14.2 years and 40.7 ± 13.5 years, respectively. Note that the patient groups were only matched in age and occupation with their respective control groups; the study was not designed to compare PVH and NPVH patient groups.

Informed consent was obtained from all the subjects participating in this study, and experimental protocols were approved by the institutional review board of Partners HealthCare System at Massachusetts General Hospital. Subjects were enrolled in a larger study on smartphone-based ambulatory voice monitoring (Mehta et al., 2015). For this study, only data from women were used due to the higher incidence of female patients with VH in the study sample, which reflects the incidence in the population (Kunduk & McWhorter, 2009) and the desire to control for sex-specific voice characteristics.

Table 2. Mean (standard deviation) of Voice-Related Quality of Life (V-RQOL) and Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) ratings for the patient groups with phonotraumatic (PVH) and nonphonotraumatic (NPVH) vocal hyperfunction.

Assessment	PVH	NPVH
V-RQOL		
Social-emotional	76.3 (21.7)	70.9 (29.1)
Physical functioning	61.4 (22.0)	65.4 (21.9)
Total score	67.5 (19.5)	67.8 (23.2)
CAPE-V		
Overall severity	34.3 (13.2)	25.4 (21.2)
Roughness	19.8 (13.1)	12.8 (10.5)
Breathiness	17.1 (14.1)	5.6 (8.6)
Strain	23.7 (14.4)	16.3 (20.9)
Pitch	9.8 (13.3)	6.8 (9.8)
Loudness	6.3 (11.0)	5.6 (9.0)

Data Acquisition Protocol

The data acquisition protocol, which was based on methods used in previous studies (Hillman et al., 1989; Holmberg, Hillman, & Perkell, 1988; Holmberg, Hillman, Perkell, & Gress, 1994), enabled the noninvasive estimation of glottal airflow (from the oral airflow), SGP (from intraoral air pressure), and acoustic measures of vocal function during phonation. Subjects were asked to produce three sets of five consecutive /pae/ syllables in two different loudness conditions (comfortable and loud). Subjects were free to choose levels that were most natural for them without any prescribed levels of absolute pitch and loudness (however, subjects were instructed to maintain a constant pitch and loudness within each syllable string). A posteriori analysis showed that the SPL of the loud condition was approximately 6 dB higher, on average, than that of the comfortable condition.

During the syllable production, simultaneous recordings were obtained of the (a) OVV using a circumferentially vented high-bandwidth pneumotachograph mask (Glottal Enterprises, Syracuse, NY) with an effective bandwidth of approximately 0 Hz to 1.2 kHz, (b) intraoral pressure (IOP) using a catheter passed between the lips and connected to a low-bandwidth pressure sensor with an effective bandwidth of approximately 0 Hz to 80 Hz, and (c) the acoustic signal using a condenser microphone (MIC; MKE104, Sennheiser Electronic GmbH, Wedemark, Germany) placed 10 cm from the lips and having a bandwidth greater than 10 kHz. Acoustic and aerodynamic signals were low-pass filtered with an 8 kHz cutoff frequency (CyberAmp Model 380, Axon Instruments, Inc.) and synchronously sampled at a rate of 20 kHz and 16-bit quantization (Digidata 1440A, Axon Instruments, Inc., Union City, CA).

OVV, IOP, and MIC signals were calibrated to physical units. The OVV signal was calibrated to units of mL/s using reference airflow levels (MCU-4 Pneumotach Calibration Unit, Glottal Enterprises). The IOP transducer was calibrated using a closed syringe system that provided reference levels of 0, 5, 10, 15, and 20 cm of water. The MIC signal was calibrated using a Cooper-Rand electrolarynx sound source that generated multiple reference tones at increasing intensity levels measured by a Class 2 sound-level meter (NL-20, RION, Tokyo, Japan) to map the uncalibrated voltage signal to units of pascal and dB SPL at 10 cm.

Data Analysis

The OVV signal was low-pass filtered at 1100 Hz with a 10th-order Chebyshev Type II filter and then decimated to

8192 Hz to simplify the IF procedure (focusing only on the first formant [F1]) and to avoid the antiresonance in the frequency response of the pneumotachograph mask at approximately 1500 Hz (Rothenberg, 1973). The IOP signal was low-pass filtered at 80 Hz with a fifth-order Butterworth filter and then decimated to 256 Hz. The MIC signal was rectified and low-pass filtered at 80 Hz with a fourth-order Butterworth filter and then decimated at 256 Hz to yield a root-mean-square (RMS) envelope (Perkell, Holmberg, & Hillman, 1991). All filtering processes were applied to the signals in both forward and reverse directions to yield zero-phase distortion and thus maintain time-alignment with the other physiological signals.

Airflow and acoustic measures were computed from the middle three syllables in each string to avoid voice initiation and termination effects, yielding a total of nine sets of measurements per loudness condition (Hillman et al., 1989; Holmberg et al., 1988, 1994). To avoid onset and offset effects for the vowels, 25% from the beginning and end of each vowel sample was discarded to yield a stable midvowel segment (see black arrows in Figure 1). These preprocessing steps yielded vocal function measures that could be compared with previous studies of subjects with normal voices and VH (Hillman et al., 1989; Holmberg et al., 2003, 1988).

An IF technique was applied to the OVV signal to cancel out the effects of the first F1 and estimate the glottal airflow from which measures were extracted to characterize the glottal volume velocity voicing source (Holmberg et al., 1988; Perkell, Hillman, & Holmberg, 1994). Most of the many IF algorithms that decompose voice source and vocal tract filter components rely on estimating the vocal tract transfer function during the closed phase of vocal fold vibration (Drugman, Alku, Alwan, & Yegnanarayana, 2014; Koc & Ciloglu, 2016). These techniques commonly suffer in the context of high fundamental frequencies (limiting the number of samples within the closed phase of each glottal cycle), nonlinear source–filter interaction, and pathological phonation (Drugman et al., 2014; Jinachitra & Smith, 2005; Koc & Ciloglu, 2016; Milenkovic, 1986; Vincent, Rosec, & Chonavel, 2007). The present study could be affected by such issues because all the subjects were women (generally having higher fundamental frequencies than male speakers), and half of them had voice disorders (pathological phonation with reduced closed phase and increased nonlinear source–filter interaction), making the detection of glottal closure instants and the application of closed-phase IF methods challenging. To counter these challenges, a single notch filter (SNF) IF technique (a conjugate pair of zeros with unity gain at direct current) was used to reduce waveform ripple (due to formant information) and produce nearly flat amplitude in the closed phase (Cheyne, 2006; Perkell et al., 1991). Although simple in nature, this method has been applied successfully in previous IF studies of both normal and pathological voice production (Hillman et al., 1989; Holmberg et al., 1988; Perkell et al., 1994).

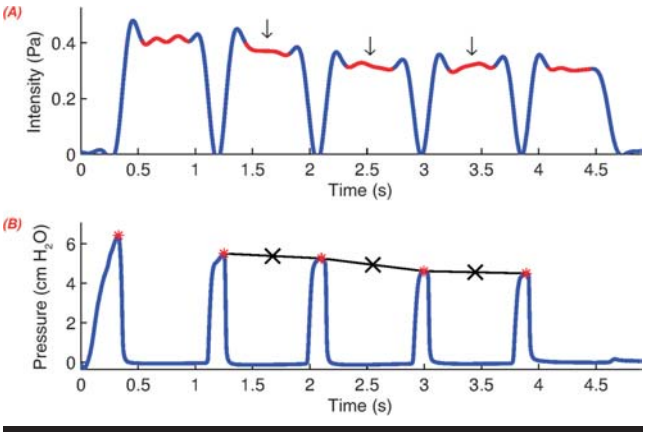
The SNF IF method has traditionally necessitated that the user make interactive expert judgments by visually assessing the IF waveform and spectrum, with the goal of minimizing formant ripple (Alku, Airas, Bäckström, & Pulakka, 2005). However, subjective user interaction is time-consuming and not suitable for analyzing large numbers of voice samples (Alku et al., 2005). Thus, we developed a simple approach to determine an initial F1 candidate in the OVV signal on the basis of minimizing the formant ripple, using the following error criterion:

$$\sum_{n=0}^{N-1} \left| \Delta^2 x_{IF}(n) \right| \tag{1}$$

where x_{IF} (n) is the inverse-filtered OVV signal at sample n, Δ^2 is the second-order time-derivative operator, and N is the number of samples. The Δ^2 operator emphasizes (12 dB/octave) the high-frequency ripple related to the F1 whose energy decreases as the center frequency of the SNF approaches F1. The SNF is applied by sweeping the center frequency from 200 Hz to 1000 Hz in 1 Hz steps. In this initial step, the filter bandwidth was fixed at 70 Hz to follow past procedures (Hillman et al., 1989; Holmberg et al., 1988; Perkell et al., 1991) and also due to the challenges of bandwidth estimation. This decision was supported by recent simulations that showed less than 20-Hz variability in the bandwidth of F1 for the /ae/ vowel (Mehta & Wolfe, 2015), which has little influence on the IF waveform.

The initial glottal airflow estimate was achieved when Equation 1 reached a minimum value, which is visually confirmed as follows. For each loudness condition, we selected one vocalic token to inverse filter for which SPL was closest to the mean SPL across the nine vocalic segments per subject. A custom MATLAB graphical user

Figure 1. Definition of low-bandwidth glottal airflow waveform measures. (A) Sound intensity (smoothed root-mean-square of the radiated acoustic pressure). Black arrows (↓) indicate midvowel segments (red line) during which glottal aerodynamic measures were computed. (B) Intraoral pressure for five /pae/ syllables showing peak values as red asterisks and interpolation lines indicating estimated subglottal pressure halfway between peaks (black Xs).



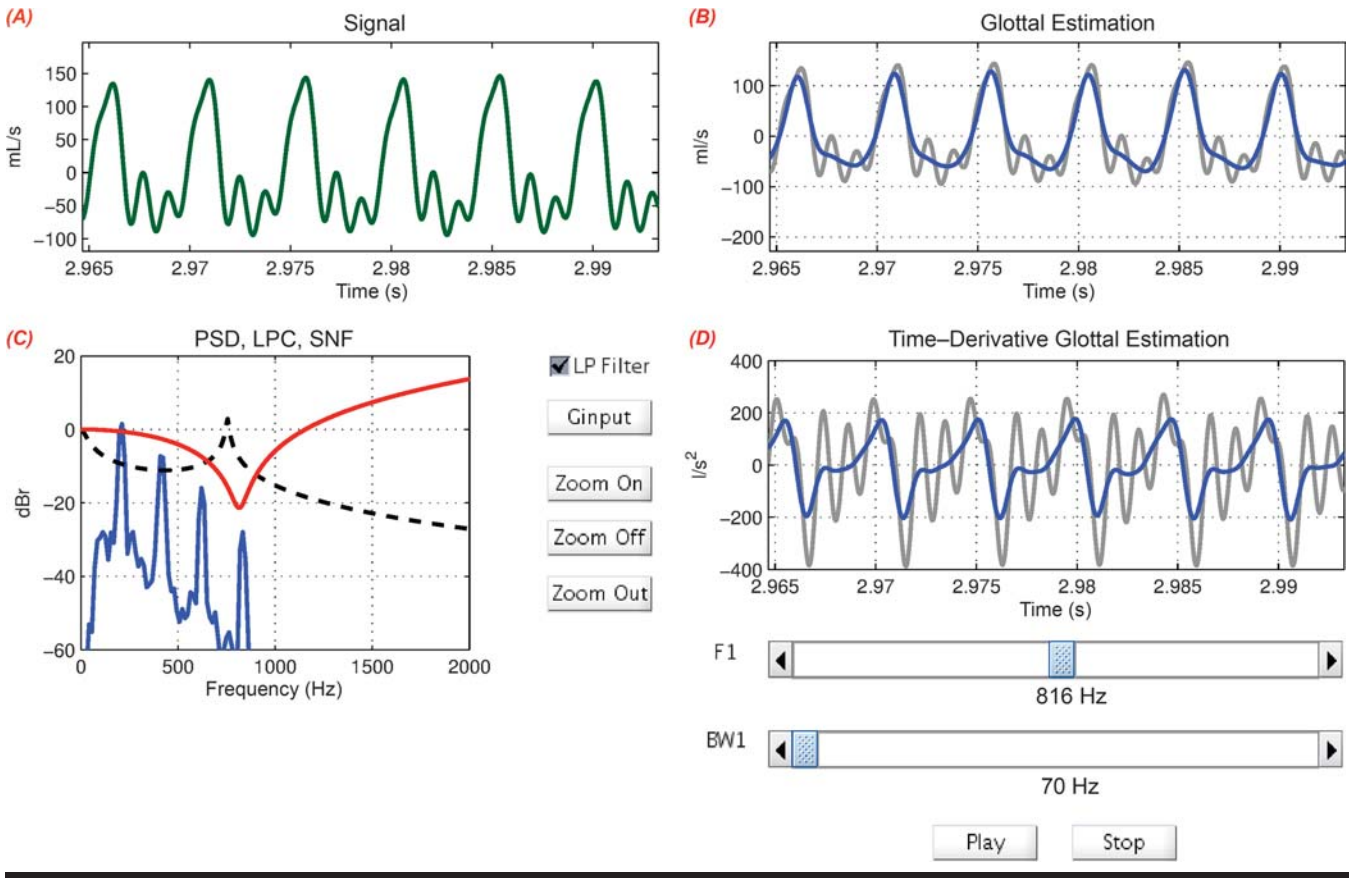
interface, shown in Figure 2, provided the ability to visually confirm and fine-tune the SNF parameters, if this was deemed necessary after the automatic process described above. Slider controls in the graphical user interface allowed the user to manually adjust the SNF center frequency (F1) and bandwidth to minimize any evidence of residual formant activity (e.g., ripple) on the basis of the visual examination of multiple displays, including (a) direct comparisons between the original OVV and resulting IF (glottal airflow) waveforms and first derivatives to check for the amount of reduction in format ripple during the closed phase, (b) spectral displays of the linear prediction–based estimates of F1 for the OVV signal and resulting notch filter using the autocorrelation method (Rabiner & Schafer, 1978) to check for evidence of residual formant energy, and (c) power density spectrum of the estimated glottal airflow signal using a Hann window of 512 samples to ensure there was a decrease in the spectral tilt.

The initial SNF parameter candidates (using the automatic approach) performed sufficiently well in approximately 70% of the cases. For the remaining data, SNF center frequencies and bandwidths were manually adjusted, with bandwidth variation having a minimal effect on the measures of interest. The linear prediction–based resonance was less efficient as a visual reference for voices with higher pitch and higher spectral tilt. Regardless, the combination of automatic and interactive approaches provided a reasonable and efficient system that also reduced the degree of uncertainty associated with the IF process, which is of particular concern when dealing with high-pitched and pathological female voices.

Measures

Low-bandwidth and high-bandwidth measures were extracted from the processed data. As shown in Figure 1, the low-bandwidth measures were taken at midvowel and included estimates of (a) average dB SPL from the RMS envelope of the acoustic signal and (b) average SGP from the average of the peak intraoral air pressures

Figure 2. Graphical user interface to aid the initial automatic IF algorithm. (A) Original oral airflow waveform without its DC component. (B) Estimated glottal airflow waveform after IF superimposed on oral airflow. (C) Power spectral density of the estimated glottal airflow waveform (PSD; solid blue), linear prediction coding spectrum (LPC; dashed black), and single notch filter frequency response (SNF; solid red). (D) Time-derivative of the estimated glottal airflow waveform superimposed on time-derivative of oral airflow waveform. Slider controls dynamically change the center frequency (F1) and bandwidth (BW1) of the inverse filter. Sound player buttons (play and stop controls) provide audio feedback to user. The Ginput (Graphical input) button is used to select a segment to perform the IF process. LP = low-pass.



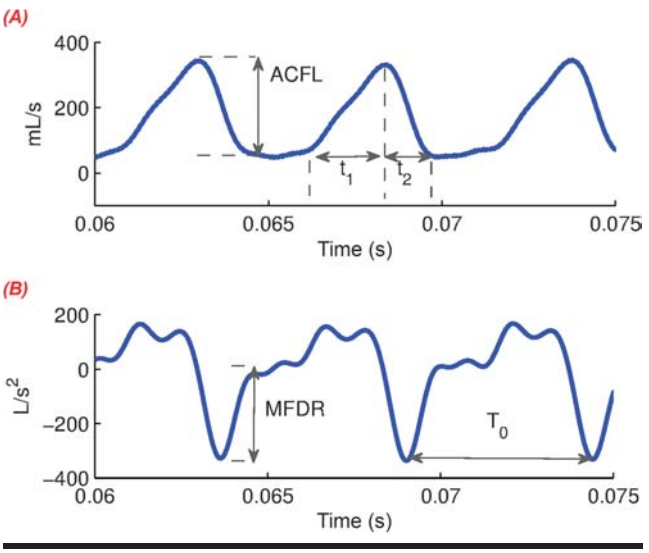
during lip closure for the /p/ sounds before and after each vowel.

As shown in Figure 3, high-bandwidth measures were taken from the IF estimates of glottal airflow and included (a) ACFL, defined as the peak-to-peak amplitude of the waveform; (b) MFDR, defined as the absolute negative peak of the first derivative of the waveform; and (c) OQ, defined as the ratio of the open phase to the total cycle duration, wherein the open and closure time points were obtained at 5% amplitude between minimum and peak flow to minimize the effect of closed-phase ripples.

Statistical Analysis

Descriptive statistics (means and standard deviations) were computed for all the measures of vocal function for each of the two loudness conditions (comfortable and loud) within each of the four subject groups. Previous work has shown that all of the aerodynamic measures in the present study are highly correlated with SPL (Holmberg et al., 1988; Holmberg, Hillman, Perkell, Guiod, & Goldman, 1995) and that adjusting/normalizing for this relationship improves the sensitivity of the measures for differentiating normal and pathological vocal function (Hillman et al., 1989). Thus, each measure was also normalized with respect to SPL (dividing SPL by a given measure) to facilitate comparisons between groups (i.e., to increase the sensitivity of the measures by controlling for the impact of variations in SPL). For SGP, ACFL, and MFDR, the normalization process entailed first converting the measured values into logarithm scales (20 times the common logarithm of each measure) before computing the ratios. Log-scaling linearized the relationship between these glottal aerodynamic measures and SPL. This step was not performed for OQ because OQ is essentially a percentage representing a time-based relationship. Note that the normalization process produced ratios that may be interpreted as larger values reflecting more "efficient" voice production (i.e., higher SPL relative to a given aerodynamic measure). In the following, note normalized parameters use prime notation (e.g., MFDR′).

Figure 3. Definition of high-bandwidth glottal airflow waveform measures. (A) Estimated glottal airflow waveform, where peak-to-peak airflow (ACFL) is defined as the peak-to-peak waveform amplitude and open quotient = $(t_1 + t_2) / T_0$, where t_1 is the opening phase duration, t_2 is the closing phase duration, and T_0 is the time interval between two consecutive peaks of the (B) time-derivative of the estimated glottal airflow waveform. Maximum flow declination rate (MFDR) is defined as the maximum negative peak in the derivative waveform.



Before applying SPL normalization, the strength of these relationships was confirmed by applying the same methods used in this study to analyze data from a larger group of 78 adult female subjects with normal voices that included the matched control subjects in the current study. The strength of relationships between each measure and SPL was evaluated by using pairwise linear regression to yield the coefficient of determination (R^2) and slope of the regression line. The results, shown in Table 3, confirmed that SPL is highly correlated with almost all of the aerodynamic measures being used in this study and thus justify the use of SPL-based normalization. OQ is the exception, but we include it because better results were found using normalization. It is also reassuring that the changes in SPL that were observed for a doubling in ACFL, MFDR, and SGP were in very close agreement with results reported in the literature (Alku, Airas, Björkner, & Sundberg, 2006; Björklund & Sundberg, 2015; Fant, 1982; Holmberg et al., 1988). In particular, Björklund and Sundberg (2015) found for women an 11.1-dB increase in SPL per doubling in SGP, which compares well with the 11.0-dB increase in SPL observed in the present study (see Table 3). Pairwise linear regression was also used to assess the relationships between f_o and each measure in the larger group of healthy women, but it did not yield strong linear correlations ($R^2 < .49$) and thus was not included in the normalization process.

Statistical testing was performed on the SPL-normalized data. Given that the two groups with voice disorders were each carefully matched to separate groups of healthy subjects, comparisons were only carried out between each patient group (PVH or NPVH) and its respective control group. Group-based comparisons were first evaluated with a multivariate t test (paired-samples Hotelling's T^2) using all features in a four-dimensional space (ACFL′, MFDR′, SGP′, and OQ′). If statistical significance was found using the multivariate t test or if associated effect size magnitudes were large (Cohen's $d > 0.6$; Cohen, 1988), follow-up comparisons were performed using one-tailed paired t tests (all hypotheses predicted larger parameter values for the disordered groups, $p < .05$) to determine the individual contribution of each feature.

Results

Table 4 reports the descriptive statistics for the original (unnormalized) measures within each subject group and loudness condition. In general, it appears that the PVH group displayed higher values across all measures

Table 3. Coefficient of determination (R^2), slope, and change in SPL per measure doubling based on pairwise linear regressions between each aerodynamic measure and SPL for a group of 78 women with normal voices.

Measure	R^2	Slope (dB/dB)	Change in SPL per doubling of measure (dB)
ACFL	0.61	1.46	8.7
MFDR	0.68	1.17	7.0
SGP	0.73	1.85	11.0
OQ	0.42	–0.40[a]	20.0

Note. Doubling is defined as +6 dB for ACFL, MFDR, and SGP. ACFL = peak-to-peak airflow; MFDR = maximum flow declination rate; SGP = subglottal air pressure.

[a]In dB/pp units.

than its control group did and that the measures for the NPVH cohort tended to be approximately equivalent or slightly lower than those obtained for its control group. Table 5 displays the descriptive statistics for the SPL-normalized measures within each subject group and loudness condition.

Table 6 summarizes results from statistical tests of the SPL-normalized data in Table 5. Overall, the PVH group displayed statistically lower SPL-normalized values than those in its control group: comfortable condition, $F(4, 11) = 6.45$, $p = .006$; loud condition, $F(4, 11) = 6.69$, $p = .006$, with large associated effect size magnitudes. Follow-up paired t tests demonstrated statistically significant differences for all of the aerodynamic measures in both loudness conditions, with large effect size magnitudes for all comparisons except for MFDR′ in the comfortable voice condition (moderate effect size magnitude of 0.53). Overall, the NPVH group displayed statistically lower SPL-normalized values than those in its control group: comfortable condition, $F(4, 11) = 3.19$, $p = .057$; loud condition, $F(4, 11) = 4.91$, $p = .008$, with large associated effect size magnitudes. Follow-up paired t tests demonstrated statistically significant differences and large effect size magnitudes for OQ′ in the comfortable and loud conditions and for SGP′ in the comfortable condition.

Discussion

This study sought to confirm preliminary evidence that glottal aerodynamic measures could identify pathophysiological mechanisms for PVH and NPVH (Hillman et al., 1989) that are each distinctly different from normal vocal function. Statistically significant results from data collected in carefully selected cohorts of patients with PVH and NPVH and well-matched control groups provide support for the observations and hypotheses that were based on this earlier work. Namely, PVH patients displayed significantly reduced SPL-normalized values for the aerodynamic measures of SGP, ACFL, MFDR, and OQ. This means that higher than normal levels of these parameters are needed to attain a given SPL (reduced "vocal efficiency"), thus reflecting increased potential for trauma to vocal fold

Table 4. Group mean (standard deviation) for aerodynamic and SPL measures from the /pae/ syllable productions in comfortable and loud voices for the PVH and NPVH patient groups and associated matched control groups with normal voices.

Measure	PVH controls	PVH group	NPVH controls	NPVH group
ACFL (mL/s)				
Comfortable	205 (63)	296 (102)	271 (94)	220 (77)
Loud	264 (90)	400 (141)	340 (123)	302 (112)
MFDR (L/s^2)				
Comfortable	306 (131)	415 (177)	386 (204)	269 (128)
Loud	418 (189)	648 (309)	573 (314)	491 (248)
SGP (cm H_2O)				
Comfortable	8.2 (1.6)	12.7 (4.5)	8.6 (2.7)	8.8 (1.6)
Loud	11.5 (1.8)	17.6 (5.2)	13.2 (3.8)	13.4 (3.4)
OQ (%)				
Comfortable	67.9 (10.7)	87.0 (8.3)	70.3 (8.6)	78.1 (10.3)
Loud	65.8 (12.8)	81.1 (10.1)	58.7 (8.6)	63.0 (7.3)
SPL (dB SPL)				
Comfortable	83.0 (5.0)	84.4 (4.6)	84.2 (5.4)	81.8 (5.9)
Loud	89.2 (4.9)	91.3 (4.6)	92.4 (4.1)	90.1 (5.3)

Note. PVH = phonotraumatic vocal hyperfunction; NPVH = nonphonotraumatic vocal hyperfunction; ACFL = peak-to-peak airflow; MFDR = maximum flow declination rate; SGP = subglottal air pressure; OQ = open quotient.

Table 5. Group mean (standard deviation) for SPL-normalized/log-scaled aerodynamic measures from the /pae/ syllable productions in comfortable and loud voice for the PVH and NPVH patient groups and associated matched control groups with normal voices.

Measure	PVH controls	PVH group	NPVH controls	NPVH group
ACFL′ (dB/dB)				
Comfortable	1.81 (0.10)	1.73 (0.10)	1.75 (0.09)	1.77 (0.08)
Loud	1.86 (0.10)	1.78 (0.11)	1.85 (0.10)	1.84 (0.10)
MFDR′ (dB/dB)				
Comfortable	1.70 (0.09)	1.65 (0.10)	1.67 (0.10)	1.72 (0.08)
Loud	1.73 (0.09)	1.66 (0.11)	1.72 (0.13)	1.72 (0.11)
SGP′ (dB/dB)				
Comfortable	4.62 (0.41)	3.96 (0.45)	4.67 (0.60)	4.38 (0.33)
Loud	4.24 (0.28)	3.75 (0.38)	4.24 (0.47)	4.05 (0.25)
OQ′ (dB/pp)				
Comfortable	1.26 (0.26)	0.98 (0.13)	1.22 (0.20)	1.07 (0.21)
Loud	1.41 (0.32)	1.13 (0.13)	1.61 (0.28)	1.45 (0.20)

Note. PVH = phonotraumatic vocal hyperfunction; NPVH = nonphonotraumatic vocal hyperfunction; ACFL = peak-to-peak airflow; MFDR = maximum flow declination rate; SGP = subglottal air pressure; OQ = open quotient; pp = percentage points.

tissue that would contribute to the chronic presence of vocal fold lesions and associated dysphonia in this group. These findings, including the relatively larger effect size magnitudes for SGP′ and ACFL′ than for MFDR′, are in agreement with the results of modeling VH reported by Zañartu et al. (2014) that demonstrated that compensatory mechanisms could account for the increases in aerodynamic measures. Specifically, increasing SGP to maintain a given SPL when there is reduced glottal closure (e.g., obstruction of glottal closure by vocal fold pathology) results in an elevation of ACFL and MFDR, with a concomitant increase in vocal fold collision forces. The combined results from the present study and modeling work (Zañartu et al., 2014) reflect the vicious cycle that is associated with PVH in which a compensatory increase in vocal effort could also cause additional vocal fold trauma (Hillman et al., 1989). In contrast, NPVH patients displayed abnormally lower SPL-normalized values for SGP and OQ only. This means that whereas higher than normal levels of these two parameters are needed to attain a given SPL (reduced "vocal efficiency"), the lack of a concomitant increase in ACFL and MFDR reflects decreased potential to cause trauma to vocal fold tissue.

Table 6. Results of between-group statistical comparisons using Table 5 data; reported are effect sizes for the multivariate, paired-samples Hotelling's T^2 tests and univariate, one-tailed paired *t* tests (Cohen's *d*).

Group comparison	Hotelling's T^2	ACFL′	MFDR′	SGP′	OQ′
PVH vs. Controls					
Comfortable	1.48***	−0.80***	−0.53**	−1.53***	−1.36***
Loud	1.51***	−0.76***	−0.70**	−1.47***	−1.11***
NPVH vs. Controls					
Comfortable	1.04*	—	—	−0.60**	−0.73**
Loud	1.29***	—	—	—	−0.66**

Note. Negative values for the univariate effect sizes signify that SPL-normalized measures are smaller in the patient groups than in their respective control groups. ACFL = peak-to-peak airflow; MFDR = maximum flow declination rate; SGP = subglottal air pressure; OQ = open quotient; PVH = phonotraumatic vocal hyperfunction; NPVH = nonphonotraumatic vocal hyperfunction.

*$p < .056$. **$p < .05$. ***$p = .025$.

The data for this study came entirely from adult female subjects because of the higher incidence of VH in women (Kunduk & McWhorter, 2009) and the desire to maintain homogeneity of the groups. However, most of the experimental (e.g., excised larynges, magnetic resonance imaging, and mechanical models) and theoretical (e.g., numerical models) studies of voice production have been based on data from male subjects. IF the F1 was therefore more challenging than reported in much of the literature because the higher fundamental frequency and spectral tilt associated with female voices increased the uncertainty of the estimation process. For example, as fundamental frequency increases, the number of harmonics is limited by the bandwidth of our measures (1 kHz). Therefore, a common female fundamental frequency of 300 Hz or higher will provide only two harmonics (600 and 900 Hz), which will have a direct influence on parameters that include high-frequency information (e.g., MFDR). Furthermore, recent results from numerical vocal fold models show that the closed phase is not completely flat (Zañartu et al., 2014); thus, even in the presence of incomplete glottal closure, the ripple-minimization approach (Lindqvist-Gauffin, 1964) may not be the best physiological criteria for IF. For signals with higher spectral tilt, we observed small differences in the resulting waveforms when varying the center frequency and bandwidth of the SNF used to cancel out the F1. The waveform for these cases follow a sinusoidal pattern that may influence the accuracy of parameters such as OQ and MFDR.

As a historical reference, Table 7 compares the SPL-normalized aerodynamic measures computed in the present study and those derived (i.e., estimated using their SPL and log-transformed parameters) from published data (Hillman et al., 1989; Holmberg et al., 1988; Sapienza & Stathopoulos, 1994). The general observation here is that the ratios in the current study are in similar numerical ranges to those derived from previously published data. Also, the estimated SPL-normalized aerodynamic measures from past studies decrease for both types of VH compared with measures in subjects with normal voices, which is in agreement with the current results.

The positive results of the present study support the potential for glottal aerodynamic measures to objectively quantify pathophysiological mechanisms for PVH and NPVH that are each distinctly different from normal vocal function. Treatment-related investigations should also be undertaken to further assess the value of the measures because clinically meaningful measures would ideally

Table 7. Mean SPL-normalized glottal aerodynamic measures from current and previously published investigations studying adult female subjects.

Measure and source	Comfortable			Loud		
	Normal	PVH	NPVH	Normal	PVH	NPVH
ACFL′ (dB/dB)						
Mean values from Table 5	1.78	1.73	1.77	1.86	1.78	1.84
Holmberg et al., 1988	1.78			1.85		
Holmberg et al., 2003		1.67			1.73	
Hillman et al., 1989			1.74			1.73
MFDR′ (dB/dB)						
Mean values from Table 5	1.69	1.65	1.72	1.73	1.66	1.72
Holmberg et al., 1988	1.72			1.74		
Holmberg et al., 2003		1.54			1.55	
Hillman et al., 1989			1.68			1.70
SGP′ (dB/dB)						
Mean values from Table 5	4.65	3.96	4.38	4.24	3.75	4.05
Holmberg et al., 1988	5.00			4.56		
Holmberg et al., 2003		3.99			3.86	
Hillman et al., 1989			4.33			4.15
OQ′ (dB/pp)						
Mean values from Table 5	1.24	0.98	1.07	1.51	1.13	1.45
Holmberg et al., 1988	1.01			1.17		
Holmberg et al., 2003		1.39			1.43	
Hillman et al., 1989			0.90			1.01

Note. Values for each measure are the means from Table 5 and means of previously reported data—specifically, 20 subjects with normal voices (Holmberg et al., 1988), 10 subjects with PVH (Holmberg et al., 2003), and two subjects with NPVH (Hillman et al., 1989); PVH = phonotraumatic vocal hyperfunction; NPVH = nonphonotraumatic vocal hyperfunction; ACFL = peak-to-peak airflow; MFDR = maximum flow declination rate; SGP = subglottal air pressure; OQ = open quotient; pp = percentage points.

migrate toward normal values after successful therapeutic intervention. For example, better insights into etiological mechanisms (on the basis of empirical evidence that is currently lacking) might be attained by using these measures to assess vocal function in PVH patients before and after surgical removal of lesions and subsequent to voice therapy. According to prevailing clinical assumptions (Leonard, 2009), it is possible that postsurgical measures would continue to reflect the persistence of VH (potential for recurrence of vocal fold trauma) and that the measures would show only significant migration toward normal after vocal retraining (voice therapy). Such expectations are somewhat tempered by previous evidence that the within-subject variability of some glottal aerodynamic measures (e.g., for repeated pretreatment baseline measures) may limit sensitivity to treatment effects (Holmberg et al., 2003). However, it is hoped that the signal analysis framework used in the current study will reduce the portion of this reported variability that may have been related to the use of older methodologies.

The recent development of subglottal impedance-based IF (Llico et al., 2015; Zañartu, Espinoza, et al., 2013; Zañartu, Ho, et al. 2013) offers the capability to extract estimates of glottal airflow waveform parameters from a neck-surface acceleration signal, thus providing the opportunity to unobtrusively obtain these promising measures as individuals go about their usual daily activities (Mehta et al., 2015). Such capabilities could provide a much more accurate assessment of an individual's typical function (e.g., the prevalence and severity of VH during a typical day) and potentially supply physiologically based biofeedback targets to aid in reducing VH (Llico et al., 2015).

Conclusion

The results of this study confirm previous hypotheses and preliminary results indicating that SPL-normalized estimates of glottal aerodynamic measures (SGP, ACFL, MFDR, and OQ) can be used to identify pathophysiological phonatory mechanisms associated with two primary manifestations of VH that are each distinctly different from normal vocal function. PVH is associated with abnormally lower values for all of the SPL-normalized glottal aerodynamic parameters, reflecting lower vocal efficiency and increased potential for trauma to vocal fold tissue. NPVH exhibits abnormally lower SPL-normalized values for SGP and OQ, but without concomitant decreases in SPL-normalized ACFL and MFDR, reflecting inefficient phonation and decreased potential for trauma to vocal fold tissue. These findings support the continued development of noninvasive glottal aerodynamic measures for clinical applications.

Acknowledgments

This research was supported in part by grants from the National Institute on Deafness and Other Communication Disorders (Grants R33 DC011588 and P50 DC015446 awarded to Robert

E. Hillman and Grant F31 DC014412 awarded to Jarrad H. Van Stan), CONICYT (Grants FONDECYT 1151077 and BASAL FB0008 awarded to Matías Zañartu), and UTFSM (Grants PIIC 2014 and FSM1204 awarded to Víctor M. Espinoza). This work was also conducted with support (statistical consultation from Dr. Mark Vangel) from Harvard Catalyst: The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102, awarded to Lee M. Nadler) and financial contributions from Harvard University and its affiliated academic health care centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health.

References

Alku, P., Airas, M., Bäckström, T., & Pulakka, H. (2005). Using group delay function to assess glottal flows estimated by inverse filtering. *Electronics Letters, 41,* 562–563.

Alku, P., Airas, M., Björkner, E., & Sundberg, J. (2006). An amplitude quotient based method to analyze changes in the shape of the glottal pulse in the regulation of vocal intensity. *The Journal of the Acoustical Society of America, 120,* 1052–1062.

Bhattacharyya, N. (2014). The prevalence of voice problems among adults in the United States. *The Laryngoscope, 124,* 2359–2362.

Björklund, S., & Sundberg, J. (2015). Relationship between subglottal pressure and sound pressure level in untrained voices. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2015.03.006

Cheyne, H. A., II. (2006). Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck. In *Proceedings of the 3rd IEEE/EMBS International Summer School on Medical Devices and Biosensors* (pp. 118–121). Boston, MA: IEEE.

Cohen, J. (1988). *Statistical power analysis for the behavior science*. Hillsdale, NJ: Erlbaum.

Drugman, T., Alku, P., Alwan, A., & Yegnanarayana, Y. (2014). Glottal source processing: From analysis to applications. *Computer Speech & Language, 28,* 1117–1138.

Fant, G. (1982). Preliminaries to analysis of the human voice source. *STL-QPSR, 4*(1982), 1–28.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32,* 373–392.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1990). Phonatory function associated with hyperfunctionally related vocal fold lesions. *Journal of Voice, 4,* 52–63.

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice, 13,* 557–569.

Holmberg, E. B., Doyle, P., Perkell, J. S., Hammarberg, B., & Hillman, R. E. (2003). Aerodynamic and acoustic voice measurements of patients with vocal nodules: Variation in baseline and changes across voice therapy. *Journal of Voice, 17,* 269–282.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal air-flow and transglottal air-pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America, 84,* 511–529.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., & Gress, C. (1994). Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings. *Journal of Speech, Language, and Hearing Research, 37,* 484–495.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research, 38,* 1212–1223.

Jinachitra, P., & Smith, J. O. (2005). Joint estimation of glottal source and vocal tract for vocal synthesis using Kalman smoothing and EM algorithm. In *Proceedings of the 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (pp. 327–330). New Paltz, NY: IEEE.

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18,* 124–132.

Koc, T., & Ciloglu, T. (2016). Nonlinear interactive source-filter models for speech. *Computer Speech & Language, 36,* 365–394.

Kunduk, M., & McWhorter, A. J. (2009). True vocal fold nodules: The role of differential diagnosis. *Current Opinion in Otolaryngology & Head and Neck Surgery, 17,* 449–452.

Leonard, R. (2009). Voice therapy and vocal nodules in adults. *Current Opinion in Otolaryngology & Head and Neck Surgery, 17,* 453–457.

Lindqvist-Gauffin, J. (1964). Inverse filtering. Instrumentation and techniques. *Speech Transmission Laboratory Quarterly Progress and Status Report, 5*(4), 1–4.

Llico, A., Zañartu, M., González, A., Wodicka, G., Mehta, D., Van Stan, J., & Hillman, R. (2015). Real-time estimation of aerodynamic features for ambulatory voice biofeedback [Express letters]. *The Journal of the Acoustical Society of America, 138,* EL14–EL19.

Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., . . . Hillman, R. E. (2015). Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in Bioengineering and Biotechnology, 3,* 155.

Mehta, D. D., & Wolfe, P. J. (2015). Statistical properties of linear prediction analysis underlying the challenge of formant bandwidth estimation. *The Journal of the Acoustical Society of America, 137,* 944–950.

Milenkovic, P. (1986). Glottal inverse filtering by joint estimation of an AR system with a linear input model. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 34,* 28–42.

Perkell, J. S., Hillman, R. E., & Holmberg, E. B. (1994). Group differences in measures of voice production and revised values of maximum airflow declination rate. *The Journal of the Acoustical Society of America, 96,* 695–698.

Perkell, J. S., Holmberg, E. B., & Hillman, R. E. (1991). A system for signal processing and data extraction from aerodynamic, acoustic, and electroglottographic signals in the study of voice production. *The Journal of the Acoustical Society of America, 89,* 1777–1781.

Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals* (U.S. ed.). Englewood Cliffs, NJ: Prentice Hall.

Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America, 53,* 1632–1645.

Sapienza, C. M., & Stathopoulos, E. T. (1994). Respiratory and laryngeal measures of children and women with bilateral vocal

fold nodules. *Journal of Speech and Hearing Research, 37,* 1229–1243.

Vincent, D., Rosec, O., & Chonavel, T. (2007). A new method for speech synthesis and transformation based on an ARX-LF source-filter decomposition and HNM modeling. In *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing* (Vol. 4, pp. IV-525–IV-528). Honolulu, HI: IEEE.

Zañartu, M., Espinoza, V. M., Mehta, D. D., Van Stan, J. H., Cheyne, H. A., II, Ghassemi, M., . . . Hillman, R. E. (2013). Toward an objective aerodynamic assessment of vocal hyperfunction using a voice health monitor. In C. Manfredi (Ed.), *8th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA 2013), December 16–18 2013, Firenze, Italy* (pp. 167–170). Firenze, Italy: Firenze University Press.

Zañartu, M., Galindo, G. E., Erath, B. D., Peterson, S. D., Wodicka, G. R., & Hillman, R. E. (2014). Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *The Journal of the Acoustical Society of America, 136,* 3262–3271.

Zañartu, M., Ho, J. C., Mehta, D. D., Hillman, R. E., & Wodicka, G. R. (2013). Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Transactions on Audio, Speech, and Language Processing, 21,* 1929–1939.