

## Research Article

# Glottal Aerodynamics Estimated From Neck-Surface Vibration in Women With Phonotraumatic and Nonphonotraumatic Vocal Hyperfunction

Víctor M. Espinoza,<sup>a</sup>  Daryush D. Mehta,<sup>b,c,d</sup>  Jarrad H. Van Stan,<sup>b,c,d</sup>  
Robert E. Hillman,<sup>b,c,d</sup> and Matías Zañartu<sup>e</sup>

**Purpose:** The purpose of this study was to determine whether estimates of glottal aerodynamic measures based on neck-surface vibration are comparable to those previously obtained using oral airflow and air pressure signals (Espinoza et al., 2017) in terms of discriminating patients with phonotraumatic and nonphonotraumatic vocal hyperfunction (PVH and NPVH) from vocally healthy controls.

**Method:** Consecutive /pae/ syllables at comfortable and loud level were produced by 16 women with PVH (organic vocal fold lesions), 16 women with NPVH (primary muscle tension dysphonia), and 32 vocally healthy women who were each matched to a patient according to age and occupation. Subglottal impedance-based inverse filtering of the anterior neck-surface accelerometer (ACC) signal yielded estimates of peak-to-peak glottal airflow, open quotient, and maximum flow declination rate. Average subglottal pressure and microphone-based sound pressure level (SPL) were also estimated from the ACC signal using subject-specific linear regression models. The ACC-based measures of glottal aerodynamics were normalized for SPL

and statistically compared between each patient and matched-control group.

**Results:** Patients with PVH and NPVH exhibited lower SPL-normalized glottal aerodynamics values than their respective control subjects ( $p$  values ranging from  $< .01$  to  $.07$ ) with very large effect sizes (1.04–2.16), regardless of loudness condition or measurement method (i.e., ACC-based values maintained discriminatory power).

**Conclusions:** The results of this study demonstrate that ACC-based estimates of most glottal aerodynamic measures are comparable to those previously obtained from oral airflow and air pressure (Espinoza et al., 2017) in terms of differentiating between hyperfunctional (PVH and NPVH) and normal vocal function. ACC-based estimates of glottal aerodynamic measures may be used to assess vocal function during continuous speech and enables this assessment of daily voice use during ambulatory monitoring to provide better insight into the pathophysiological mechanisms associated with vocal hyperfunction.

<sup>a</sup>Department of Sound, Universidad de Chile, Santiago

<sup>b</sup>Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston

<sup>c</sup>Speech and Hearing Bioscience and Technology Program, Harvard Medical School, Boston, MA

<sup>d</sup>MGH Institute of Health Professions, Boston, MA

<sup>e</sup>Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

Correspondence to Matías Zañartu: [matias.zanartu@usm.cl](mailto:matias.zanartu@usm.cl)

Editor-in-Chief: Bharath Chandrasekaran

Editor: Jack J. Jiang

Received April 20, 2020

Revision received May 29, 2020

Accepted June 10, 2020

[https://doi.org/10.1044/2020\\_JSLHR-20-00189](https://doi.org/10.1044/2020_JSLHR-20-00189)

Voice disorders affect approximately 30% of the adult population in the United States at some point in their lives (Roy et al., 2005), with the most common being associated with vocal hyperfunction (VH; Bhattacharyya, 2014). VH refers to “chronic conditions of abuse and/or misuse of the vocal mechanism due to excessive and/or ‘imbalanced’ muscular forces” (Hillman et al., 1989) and manifests as two types of disorders: (a) phonotraumatic VH (PVH), which causes trauma to vocal fold tissue and the formation of lesions (e.g., nodules and polyps); and (b) non-phonotraumatic VH (NPVH), which causes dysphonia and vocal fatigue in the absence of vocal fold tissue trauma or other conditions that could affect phonation (Mehta et al., 2015).

**Disclosure:** Robert Hillman has a financial interest in InnoVoice LLC, a company focused on developing and commercializing technologies for the prevention, diagnosis and treatment of voice-related disorders. Robert Hillman’s interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies.

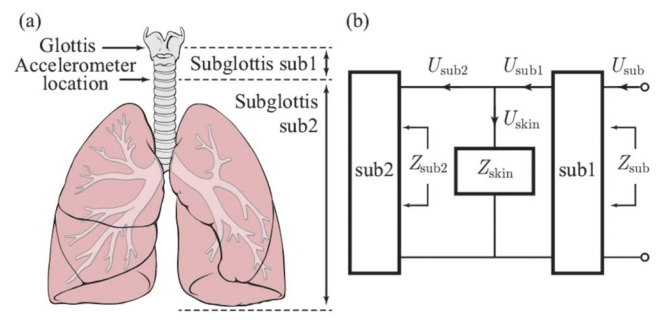
Previous work has shown that the pathophysiological mechanisms for the two types of VH can be described and differentiated using estimates of glottal aerodynamic parameters that are normalized for sound pressure level (SPL; Espinoza et al., 2017; Hillman et al., 1989). The measures are obtained in the laboratory using specialized equipment and methods that include inverse filtering the oral volume velocity (OVV) airflow that is captured using a circumferentially vented face mask to achieve an adequate frequency response (Rothenberg, 1973). The salient measures are extracted from acoustic and aerodynamic recordings of /pae/ nonsense syllable repetitions and include subglottal air pressure (SGP, estimated from the intraoral air pressure [IOP] during the /p/ consonants) and three measures extracted from inverse-filtered estimates of the glottal volume velocity waveform during the /ae/ vowels: (a) peak-to-peak amplitude of the unsteady airflow (ACFL), (b) maximum flow declination rate (MFDR, defined as the absolute negative peak of the first derivative of the waveform), and (c) open quotient (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).

A recent study (Espinoza et al., 2017) used the SPL-normalized aerodynamic measures to assess vocal function in women diagnosed as having PVH or NPVH along with matched vocally healthy controls that were in large enough groups to statistically validate the observations made in an earlier descriptive study of VH (Hillman et al., 1989). The statistical results from the Espinoza et al., (2017) study are summarized in Table 3 along with the results of this study to facilitate comparisons. Group-based comparisons using all of the aerodynamic measures showed statistically significant differences between each of the voice disordered groups (PVH and NPVH) and their respective control groups. In both the Hillman et al. (1989) and Espinoza et al. (2017) studies, the PVH patients used higher than normal levels of all of the aerodynamic parameters to attain a given SPL in comfortable and loud voice conditions (indicated by lower SPL-normalized values), which was interpreted as “reflecting increased potential for trauma to vocal fold tissue that would contribute to the chronic presence of vocal fold lesions and associated dysphonia in this group” (Espinoza et al., 2017, p. 2166). This interpretation is further supported by studies that used computer modeling of VH (Galindo et al., 2017; Zañartu et al., 2014) to demonstrate that 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 model results are viewed as reflecting 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 to the results for PVH patients, NPVH patients only displayed abnormally increased values for SGP and OQ (indicated by lower SPL-normalized values). These results were interpreted to mean “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” (Espinoza et al., 2017, p. 2166).

The cause and impact of hyperfunctional voice disorders are believed to be strongly associated with daily voice use (Hillman et al., 1989) that cannot be adequately characterized/assessed during a brief laboratory or clinical voice evaluation. Recent studies have demonstrated the valuable new information that can be obtained about these disorders from ambulatory voice monitoring that uses a neck-placed accelerometer (ACC) as the phonation sensor (Castellana et al., 2018; Cortés et al., 2018; Ghassemi et al., 2014; Lei et al., 2019; Van Stan et al., 2020). For example, data from weeklong monitoring are beginning to reveal differences between the vocal behavior of patients with PVH and vocally healthy matched controls using distributional (nonaverage) characteristics of SPL, fundamental frequency, and spectral tilt (Van Stan et al., 2020). In addition to these acoustically related voice measures, ACC-based estimates of glottal aerodynamics during daily life would provide important complementary features that can be physiologically interpreted in terms of hyperfunctional vocal behaviors. Fryd et al. (2016) showed that SGP can be estimated from the magnitude of the ACC signal. Zañartu et al. (2013) developed a method for estimating the glottal airflow waveform from the ACC signal that is referred to as subglottal impedance-based inverse filtering (IBIF; see Figure 1) to enable the ACC-based estimation of ACFL, MFDR, and OQ. It has already been demonstrated that IBIF can be implemented in real time on the type of smartphone that has been used as the data collection and processing platform for ambulatory voice monitoring to potentially be used for biofeedback (Castellana et al., 2018; Lei et al., 2019; Llico et al., 2015). The capability to unobtrusively obtain glottal aerodynamic measures during natural connected speech (instead of being restricted to /pae/ syllables) and during activities of daily living (ambulatory monitoring) would greatly expand the potential to more fully characterize the pathophysiology and daily impact of hyperfunctional

**Figure 1.** Representation of the neck skin and subglottal system. (a) Accelerometer position and *sub1* and *sub2* system parts. (b) A mechano-acoustic analogy of the subglottal system including load impedance from skin. © 2013 IEEE. Reprinted, with permission, from Zañartu et al. (2013).



voice disorders and lead to the development of more ecologically sound clinical voice assessment measures, as well as expand options for parameters that can be used for ambulatory biofeedback to treat VH (Van Stan et al., 2017).

In spite of the progress that has been made, the use of IBIF to extract estimates of glottal aerodynamic measures from the ACC signal has never been fully validated for disordered voices through direct comparison of the same parameters obtained from traditional oral measurements (OVV and IOP). Thus, the purpose of this study was to determine whether ACC-based estimates of glottal aerodynamic measures are comparable to those previously reported by Espinoza et al. (2017) in terms of discriminating between healthy controls and patients with PVH or NPVH.

## Method

### Participants

All participants were the same adult women as those used in the Espinoza et al. (2017) study. To briefly summarize, the disordered groups were comprised of 16 patients with PVH (vocal fold nodules or polyps) and 16 patients with NPVH (primary muscle tension dysphonia). Diagnoses were based on a complete team evaluation by laryngologists and speech-language pathologists at the Massachusetts General Hospital Voice Center that included endoscopic imaging of the larynx (Mehta & Hillman, 2012) and a complete battery of perceptual (Consensus Auditory-Perceptual Evaluation of Voice, CAPE-V; Kempster et al., 2009), instrumental (Patel et al., 2018), and patient self-assessment (Voice-Related Quality of Life [V-RQOL]; Hogikyan & Sethuraman, 1999) measures. All patients were enrolled prior to receiving any treatment. Each patient was matched with a normal control participant based on sex, occupation, and age ( $\pm 5$  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 of participants (mean  $\pm$  standard deviation [SD]) was  $32 \pm 13$  years for the PVH and their matched control group, and  $40 \pm 14$  years for the NPVH and their matched control group. Mean (SD) total V-RQOL scores were 67.5 (19.5) and 67.8 (23.2) for the PVH and NPVH groups, respectively. Mean (SD) CAPE-V ratings for overall severity of dysphonia were 34.3 (13.2) and 25.4 (21.2) for the PVH and NPVH groups, respectively. Additional details about patient demographics and scores for V-RQOL and CAPE-V subscales are reported in Table 2 of Espinoza et al. (2017). Informed consent was obtained from all the participants in this study, and experimental and clinical protocols were approved by the institutional review board of Partners HealthCare System at Massachusetts General Hospital.

### Data Acquisition Protocol

Simultaneous noninvasive recordings of vocal function were obtained from (a) the acoustic signal using a condenser

microphone (MIC; MKE104, Sennheiser, Electronic GmbH) placed 10 cm from the lips and having full bandwidth (greater than 6 kHz in most of the cases), (b) OVV airflow using a circumferentially vented pneumotachograph mask (Glottal Enterprises) with a sufficient bandwidth of approximately 1.1 kHz, (c) IOP using a catheter passed between the lips and connected to a low-bandwidth pressure sensor (with a bandwidth of approximately 80 Hz), and (d) a one-axis ACC sensor placed on the anterior neck surface halfway between the thyroid prominence and the suprasternal notch. All signals were sampled at 20 kHz/16 bits (Digidata 1440A, Axon Instruments, Inc.), low-pass filtered at 8-kHz cutoff frequency (CyberAmp Model 380, Axon Instruments, Inc.), and calibrated to physical units following methods presented in Espinoza et al. (2017). A clinically certified speech-language pathologist instructed each subject to produce strings of /pae/ syllables on one breath, while holding pitch and loudness constant at a comfortable and loud (approximately 6-dB increase) voice. All recording sessions were conducted in a sound-treated room.

### Data Processing

Both ACC and OVV signals were decimated to 8192 Hz and low-pass filtered at 1100 Hz (10th-order Chebyshev Type II filter). The DC and very-low-frequency components were removed below 60 Hz (fourth-order Butterworth filter) prior to analysis. A low-pass filter was applied to the IOP signal at 80 Hz (fifth-order Butterworth filter) and then decimated to a 256 Hz sample rate (Espinoza et al., 2017; Perkell et al., 1991). The full-bandwidth MIC signal was not filtered. All the filtering processes were applied to obtain zero-phase distortion signals to avoid time-offset with the other signals. To obtain OVV-based glottal airflow measures, tokens closest to the mean SPL value for each loudness conditions (i.e., same criterion than in Espinoza et al., 2017) were selected for further analysis and processing. Inverse filtering of the OVV signal was accomplished following the same methods presented in (Espinoza et al., 2017). Estimates of SGP (Rothenberg, 1973), SPL, and skin acceleration level (SAL) measures were derived from the IOP, MIC, and ACC signals, respectively.

ACC-based measures need a calibration step to use the IBIF model in order to obtain approximations of glottal airflow from the neck-skin ACC signal. The calibration step is reported in Zañartu et al. (2013) and briefly described as follows. Subject-specific Q parameters for the IBIF model were determined to minimize the waveform error between the OVV-based glottal airflow (reference signal) and the inverse-filtered neck-skin ACC signal (signal to be matched to the reference signal), according to Equation (1):

$$e = 0.5 \times \frac{\sum_{n=0}^{N-1} |u_g(n) - \hat{u}_g(n)|}{\sum_{n=0}^{N-1} |u_g(n)|} + 0.5 \times \frac{\sum_{n=1}^{N-1} |\Delta u_g(n) - \Delta \hat{u}_g(n)|}{\sum_{n=1}^{N-1} |\Delta u_g(n)|}, \quad (1)$$

where  $N$  is the number of voice samples,  $u_g(n)$  is the OVV-based glottal airflow waveform at sample  $n$ ,  $\hat{u}_g(n)$  is the ACC-based glottal airflow estimate at sample  $n$ , and  $\Delta$  is the discrete-time time-derivative operator, that is,  $\Delta u_g(n) = u_g(n) - u_g(n-1)$ . To obtain Q parameters, a particle swarm optimization algorithm (Kennedy & Eberhart, 1995) runs 10 times per token to select (one of 10 trials) the Q set with minimum waveform error, in addition to a visual assessment of the inverse-filtered signals (see example in Figure 2) using a custom MATLAB graphical user interface (Espinoza et al., 2017). After the IBIF process, approximations of the glottal airflow from both OVV and ACC signals were used to obtain aerodynamic measures, including ACFL, MFDR, and OQ (see Figure 2 for a visual reference illustrating the ACC- and OVV-based measures).

A subject-specific linear regression model was determined for predicting ACC-based  $SGP_L$  (i.e.,  $SGP_L = 20 \cdot \log_{10} SGP$ ) from SAL using Equation (2):

$$SGP_L = \beta_0 + \beta_1 \cdot SAL, \quad (2)$$

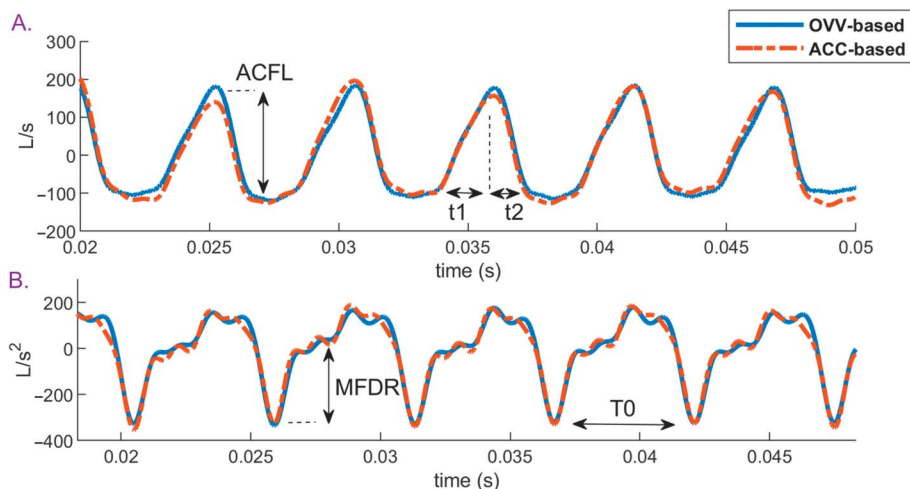
where  $SAL = 20 \cdot \log_{10}(acc_{RMS})$ ,  $(\beta_0, \beta_1)$  the model parameters, and  $acc_{RMS}$  is the root-mean-square value ( $\text{cm/s}^2$ ) of the ACC signal. This approach is based on previous work that found a strong relationship between ACC signal magnitude and IOP-derived SGP (Fryd et al., 2016; Lin et al., 2019; McKenna et al., 2017). In the current work, we are interested in the log-transformed measure of SGP that will be normalized by SPL (Espinoza et al., 2017). A similar regression model was applied to estimate SPL from SAL as well (Švec et al., 2005). An example of these models is shown in Figure 3. The SPL versus SAL and  $SGP_L$  versus SAL models were calculated for each subject, including

controls and patient groups (i.e., 64 models, 32 for the control groups and 32 for the patient groups). Then, SPL and  $SGP_L$  were determined from SAL measures using these models. ACC-based SPL-normalized measures included (labeled using prime notation) normalized peak-to-peak airflow  $ACFL' = \frac{SPL}{20 \log_{10}(ACFL)}$ , normalized maximum declination rate  $MFDR' = \frac{SPL}{20 \log_{10}(MFDR)}$ , normalized subglottal pressure  $SGP' = \frac{SPL}{20 \log_{10}(SGP)}$ , and normalized open quotient  $OQ' = \frac{SPL}{OQ}$ . 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).

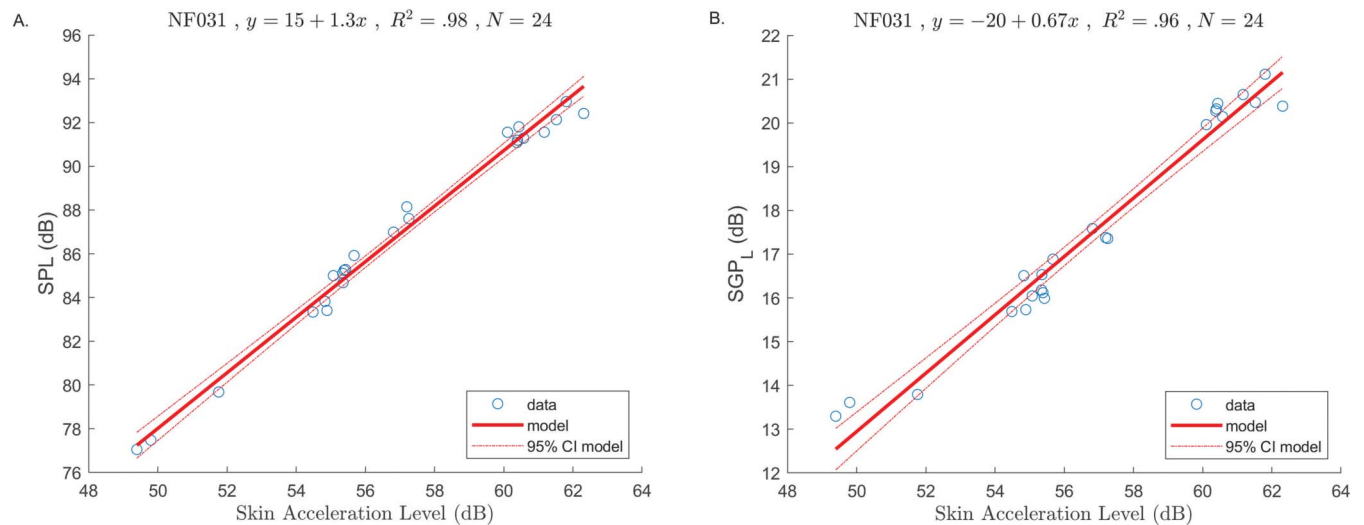
Statistical analyses followed those used by Espinoza et al. (2017). This included the calculation of descriptive statistics (mean and SD), along with multivariate and post hoc testing, to determine the extent to which the SPL-normalized ACC-based measures discriminate (i.e., show significant differences) between healthy controls and patients with PVH or NPVH. SPL normalization was done two ways: using SPL measured from the MIC (acoustic) signal and using SPL estimated from the ACC signal.

In order to validate our statistical assumptions, the normality of the data was checked using a Kolmogorov–Smirnov test (Massey, 1951), resulting in that all the data were normally distributed with  $p < .05$ . Paired multivariate Hotelling’s  $T^2$  test (one-tailed) were performed using the SPL-normalized, ACC-based measures. If overall statistically significant differences between controls and patient groups were found, a post hoc  $t$  test (one-tailed) and univariate effect sizes Cohen’s  $d$  (Cohen, 1988) were calculated to determine the strength of the differences for each SPL-normalized ACC-based measure—this discriminatory analysis was done

**Figure 2.** Definition of high-bandwidth glottal airflow waveform measures. OVV- and ACC-based glottal waveforms are shown. (A) Estimated glottal airflow waveform, indicating the peak-to-peak airflow (ACFL) as the peak-to-peak waveform amplitude and open quotient =  $100(t_1 + t_2) / T_0$ , 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. The definition of maximum flow declination rate (MFDR) is the maximum negative peak in the derivative waveform. OVV = oral volume velocity; ACC = accelerometer.



**Figure 3.** Examples of subject-specific models (SAL as predictor, SPL, and SGP<sub>L</sub> as responses). At top of each figure: Subject ID, model equation, coefficient of determination ( $R^2$ ), and sample size (N). (A) SPL versus SAL model. (B) SGP<sub>L</sub> versus SAL model. SPL = sound pressure level; SAL = skin acceleration level; SGP = subglottal air pressure; CI = confidence interval.



separately for the comfortable and loud voice conditions and for each of the two SPL normalization methods (MIC-based and SAL-based). The ACC-based results were compared with the results based on oral measurements (OVV and IOP) reported by Espinoza et al. (2017) for the same subject cohort.

## Results

Table 1 provides a summary of the descriptive statistics for the nonnormalized, ACC-based measures. Overall, the results of this study are comparable in magnitude to the reference OVV/IOP-based measures. Table 2 provides a detailed analysis of the mean bias error and associated mean absolute error of the ACC-based approach. The largest mean bias error values were observed in the comfortable loudness condition for MFDR for both PVH group, with  $-53 \text{ L/s}^2$  ( $-13\%$  error) and RMSE  $72 \text{ L/s}^2$ , followed for SGP, with  $-1.7 \text{ cm H}_2\text{O}$ , ( $-13\%$  error) and RMSE  $1.9 \text{ cm H}_2\text{O}$ . Table 3 provides a summary of the statistical test results and comparisons between previous OVV/IOP-based results (Espinoza et al., 2017) and those obtained in the current study for measures that were SPL-normalized using the MIC (labeled ACC<sup>1</sup>) and the SAL (labeled ACC<sup>2</sup>). Only values with effect sizes greater than 0.5 in magnitude are reported for both Hotelling's  $T^2$  and post hoc  $t$  tests. Negative effect sizes indicate that the patient group had lower values for the SPL-normalized measurements, which generally indicates reduced voice efficiency (i.e., the patients need to produce higher than normal glottal aerodynamic forces to attain the same values for vocal SPL as the normal group).

The results of the multivariate Hotelling's  $T^2$  analyses for the combined set of measures showed that the PVH and NPVH groups had significantly lower SPL-normalized values than those in their respective control groups ( $p$  values ranged from  $< .01$  to borderline =  $.07$ ) with very large effect sizes (1.04–2.16), regardless of voice condition (comfortable or loud), measurement method (OVV/IOP or ACC), or SPL normalization (MIC or SAL) approach that was used. Post hoc testing for significant differences in individual measures between the PVH group and their matched controls showed that the PVH patients had significantly lower SPL-normalized values for ACFL, SGP, and OQ ( $p$  values ranging from  $< .01$  to  $< .05$ ) with medium-to-large effect size magnitudes (0.66–1.68), regardless of voice condition (comfortable or loud), measurement method (OVV/IOP or ACC), or SPL normalization (MIC or SAL) approach that was used. However, the statistically significant differences ( $p < .05$ ) and medium effect sizes for OVV-based MFDR in comfortable ( $d = -0.53$ ) and loud ( $d = -0.70$ ) voice (Espinoza et al., 2017) were not seen for ACC-based MFDR in this study.

Post hoc testing for statistically significant differences in individual measures between the NPVH group and their matched controls showed that the NPVH patients had significantly lower SPL-normalized values for OQ ( $p$  values ranging from  $< .01$  to  $< .05$ ) with medium-to-large effect size magnitudes (0.66–1.01), regardless of voice condition (comfortable or loud), measurement method (OVV/IOP or ACC), or SPL normalization (MIC or SAL) approach that was used. The only other significant differences were for SGP where both sets of ACC-based measures displayed significantly lower values ( $p$  values ranging from  $< .025$  to  $.068$ ) and medium effect size magnitudes (0.62–0.69) for

**Table 1.** Group mean (standard deviation) for aerodynamic and sound pressure level (SPL) measures from the /pae/ syllable productions in comfortable and loud voice.

Measure	PVH controls		PVH group		NPVH controls		NPVH group	
	Reference	ACC	Reference	ACC	Reference	ACC	Reference	ACC
ACFL (ml/s)								
Comfortable	205 (63)	212 (69)	296 (102)	303 (113)	271 (94)	277 (108)	220 (77)	230 (82)
Loud	264 (90)	268 (86)	400 (141)	445 (163)	340 (123)	346 (149)	302 (112)	311 (121)
MFDR (L/s <sup>2</sup> )								
Comfortable	306 (131)	286 (126)	415 (177)	362 (171)	386 (204)	368 (199)	269 (128)	254 (110)
Loud	418 (189)	393 (170)	648 (309)	638 (318)	573 (314)	531 (314)	491 (248)	448 (228)
SGP (cm H <sub>2</sub> O)								
Comfortable	8.2 (1.6)	7.8 (1.7)	12.7 (4.5)	12.5 (4.5)	8.6 (2.7)	8.5 (2.6)	8.8 (1.6)	9.4 (2.8)
Loud	11.5 (1.8)	10.1 (1.7)	17.6 (5.2)	16.2 (5.2)	13.2 (3.8)	11.5 (3.5)	13.4 (3.4)	12.0 (3.0)
OQ (%)								
Comfortable	67.9 (10.7)	71.4 (11.0)	87.0 (8.3)	87.3 (8.1)	70.3 (8.6)	68.5 (10.9)	78.1 (10.3)	78.7 (12.0)
Loud	65.8 (12.8)	65.9 (11.1)	81.1 (10.1)	80.2 (9.9)	58.7 (8.6)	62.1 (13.2)	63.0 (7.3)	68.0 (7.6)
SPL (dB SPL)								
Comfortable	83.0 (5.0)	84.9 (5.0)	84.4 (4.6)	86.8 (4.6)	84.2 (5.4)	86.7 (4.9)	81.8 (5.9)	85.3 (7.8)
Loud	89.2 (4.9)	90.1 (5.2)	91.3 (4.6)	92.9 (5.9)	92.4 (4.1)	93.4 (4.4)	90.1 (5.3)	91.2 (6.9)

Note. Results are shown for the PVH and NPVH patient groups and associated matched control groups with normal voices for both reference glottal aerodynamic measures (Espinoza et al., 2017) and accelerometer-based (ACC) estimates of these measures (this study). 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.

both voice conditions, whereas the IOP-based measure of SGP was only significantly reduced ( $p < .05$ ) in the comfortable voice condition with a medium effect size ( $d = -0.60$ ; Espinoza et al., 2017).

## Discussion

The purpose of this study was to determine whether ACC-based estimates of glottal aerodynamic measures are comparable to those previously reported by Espinoza

et al. (2017) in terms of how well the ACC-based measures discriminate between healthy controls and patients with PVH or NPVH. Overall, the results show that the combination of using IBIF to extract estimates of ACFL, MFDR, and OQ from the ACC signal, and regression models to estimate SGP based on the SAL of the ACC signal, enables discrimination between matched control groups and associated PVH or NPVH groups that is comparable to that reported by Espinoza et al. (2017) in which reference measurements were obtained from the OVV and IOP signals.

**Table 2.** Mean bias error (MBE) and root mean square error (RMSE) between accelerometer-based estimates and reference measures of voice acoustics and glottal aerodynamics 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 ( $n = 16$  in each group) according to values from Table 1.

Measure	PVH controls		PVH group		NPVH controls		NPVH group	
	MBE (%)	RMSE	MBE (%)	RMSE	MBE (%)	RMSE	MBE (%)	RMSE
ACFL (ml/s)								
Comfortable	7 (3%)	14	7 (2%)	20	6 (2%)	23	10 (5%)	39
Loud	4 (2%)	10	45 (11%)	108	6 (2%)	44	9 (3%)	23
MFDR (L/s <sup>2</sup> )								
Comfortable	-20 (-7%)	26	-53 (-13%)	72	-18 (-5%)	25	-15 (-6%)	101
Loud	-25 (-6%)	34	-10 (-1%)	228	-42 (-7%)	53	-43 (-9%)	68
SGP (cm H <sub>2</sub> O)								
Comfortable	-0.4 (-4%)	0.9	-0.2 (-2%)	1.1	-0.1 (-2%)	0.7	0.6 (6%)	2.2
Loud	-1.4 (-12%)	1.5	-1.4 (-8%)	2.2	-1.7 (-13%)	1.9	-1.4 (-10%)	2.1
OQ (%)								
Comfortable	3.5 pp (5%)	5	0.3 pp (0%)	5.6	-1.8 pp (-3%)	4.8	0.6 pp (1%)	11.6
Loud	0.1 pp (0%)	6.7	-0.9 pp (-1%)	8.2	3.4 pp (6%)	8.6	5 pp (8%)	8.3
SPL (dB SPL)								
Comfortable	1.9 dB (2%)	2.5	2.4 dB (3%)	2.6	2.5 dB (-3%)	3.1	3.5 dB (4%)	7.1
Loud	0.9 dB (1%)	1.6	1.6 dB (2%)	2.5	1.0 dB (-1%)	2.0	1.1 dB (1%)	2.6

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; SPL = sound pressure level.

**Table 3.** Results of between-group statistical comparisons SPL-normalized features using the accelerometer (ACC) signal.

Group comparison	Hotelling's $T^2$			ACFL'			MFDR'			SGP'			OQ'		
	OVV	ACC <sup>1</sup>	ACC <sup>2</sup>	OVV	ACC <sup>1</sup>	ACC <sup>2</sup>	OVV	ACC <sup>1</sup>	ACC <sup>2</sup>	IOP	ACC <sup>1</sup>	ACC <sup>2</sup>	OVV	ACC <sup>1</sup>	ACC <sup>2</sup>
PVH vs. Controls															
Comfortable	1.48*	2.11 <sup>†</sup>	2.16 <sup>†</sup>	-0.80*	-0.68 <sup>+</sup>	-0.66 <sup>+</sup>	-0.53 <sup>+</sup>	-	-	-1.53*	-1.63 <sup>†</sup>	-1.68 <sup>†</sup>	-1.36*	-1.13 <sup>†</sup>	-1.08 <sup>†</sup>
Loud	1.51*	1.54 <sup>+</sup>	1.57 <sup>+</sup>	-0.76*	-1.01*	-0.89*	-0.70 <sup>+</sup>	-	-	-1.47*	-1.65 <sup>†</sup>	-1.67 <sup>†</sup>	-1.11*	-1.15 <sup>†</sup>	-1.11 <sup>†</sup>
NPVH vs. Controls															
Comfortable	1.04 <sup>a</sup>	1.62*	1.47 <sup>+</sup>	-	-	-	-	-	-	-0.60 <sup>+</sup>	-0.62*	-0.67 <sup>+</sup>	-0.73 <sup>+</sup>	-0.98*	-0.71 <sup>+</sup>
Loud	1.29*	1.45 <sup>+</sup>	1.27 <sup>+</sup>	-	-	-	-	-	-	-0.69 <sup>a</sup>	-0.60 <sup>‡</sup>	-0.66 <sup>+</sup>	-1.01 <sup>†</sup>	-0.78*	

Note. In ACC<sup>1</sup>, SPL was estimated from the microphone signal. In ACC<sup>2</sup>, SPL was estimated from the ACC signal. The OVV columns show the results from (Espinoza et al., 2017). Reported are effect sizes for the multivariate, paired-samples Hotelling's  $T^2$  tests and univariate, one-tailed paired  $t$  tests (Cohen's  $d$ ). Negative values for the univariate effect sizes signify that SPL-normalized measures are smaller in the patient groups than in their respective control groups. SPL = sound pressure level; ACFL' = SPL-normalized peak-to-peak airflow; MFDR' = SPL-normalized maximum flow declination rate; SGP' = SPL-normalized subglottal air pressure; OQ' = SPL-normalized open quotient; OVV = oral airflow volume velocity; IOP = intraoral pressure; PVH = phonotraumatic vocal hyperfunction; NPVH = nonphonotraumatic vocal hyperfunction.

<sup>†</sup> $p < .01$ . \* $p < .025$ . <sup>+</sup> $p < .05$ . <sup>a</sup> $p < .06$ . <sup>‡</sup> $p = .068$ .

The results of the current study represent the first direct validation of the use of ACC-based estimates of glottal aerodynamic measures to assess the pathophysiology of the most commonly occurring types of voice disorders (i.e., those related to VH). Since the majority of patients in the current study had only mild-to-moderate dysphonia (mean [SD] CAPE-V scores for the PVH and NPVH groups of 34.3 [13.2] and 25.4 [21.2], respectively), it is reasonable to assume that the ACC-based measures would also be sensitive enough for use in assessing and providing insights into the pathophysiology of other types of voice disorders that can often have a more severe impact on vocal function (e.g., vocal fold paralysis, glottic cancer, etc.). Being able to extract valid estimates of glottal aerodynamic measures from the neck-placed ACC also opens up the possibility of acquiring such measures in an unobtrusive way during natural connected speech, including during activities of daily living, which greatly expands the capabilities of ambulatory voice monitors that use an ACC as the phonation sensor. Expanded capabilities for ambulatory monitoring of vocal function has the potential to provide better insight into the pathophysiological mechanisms associated with voice disorders, particularly for disorders such as those associated with VH in which daily voice use is assumed to play a role. Applications to running speech obviously require further development and testing/validation, but an early attempt to use the ACC-based measures to differentiate the vocal function of phonotraumatic patients and controls based on week-long ambulatory recordings yielded promising results (Cortés et al., 2018).

It is important to point out that the effect sizes for the multivariate statistical tests of the combined measures are generally larger for the ACC-based estimates of the glottal airflow parameters than for the reference OVV-based estimates. This observation would seem to indicate that the ACC-based estimates of glottal airflow measures tend to be more sensitive to the presence of hyperfunctional voice disorders than the measures based on inverse filtering the

OVV waveform alone. Larger discriminatory power was also the case for the differences in the individual measures of SGP', where again all of the ACC-based differences were larger than the IOP-based measures, indicating the potential for greater sensitivity. This behavior may be explained by the smoothing effect introduced by the regression models for both SGP and SPL (see Figure 3). The most obvious difference among the individual measures between the ACC-based estimates and the OVV-based estimates is for MFDR'. None of the differences in ACC-based estimates of MFDR' were deemed significant, which is in contrast to the significant differences ( $p < .05$ ) found in the OVV-based estimates of MFDR' in Espinoza et al. (2017) for the comfortable (medium effect size of  $d = -0.53$ ) and loud (medium effect size of  $d = -0.70$ ) voice conditions.

Various factors can play a role when comparing the results against OVV/IOP-based estimates of glottal aerodynamic measures in Espinoza et al. (2017). For instance, the MIC signal includes turbulent aerodynamic noise occurring at the glottal level (Stevens, 2000), whereas the ACC signal has less of this component. Another possible influence in our experiments is the bandwidth of the signals. SPL from the MIC signal is calculated using full bandwidth, whereas the ACC signal has a smaller bandwidth with less high-frequency components. In addition, differences in the frequency response between the circumferentially vented mask and ACC sensor may have an impact on the inverse filtering methods, for which direct comparisons are only valid below ~1 kHz. Evidence that SGP versus SAL could be biased for healthy speakers performing nonmodal voice quality was provided in a recent study (Marks et al., 2019). For the NPVH, voice quality is likely to be nonmodal, and thus, models could be biased and with different variance. Finally, even though the ACC-based results presented here are encouraging, caution must be taken to apply these methods in other contexts. The extension of this study for different vocal gestures (e.g., running speech) is still an

open question for the methods and results we propose. For instance, even though the SPL values from SAL are in good agreement with the literature (Švec et al., 2005), excess of bias and variance could occur. SAL versus SPL models are very simple and may need to be improved (e.g., with additional predictors) for other vocal gestures.

## Conclusion

The results of this study demonstrate that laboratory estimates of most glottal aerodynamic measures based on the neck-surface acceleration signal are comparable to those previously obtained from oral airflow and air pressure signals (Espinoza et al., 2017) in terms of differentiating between hyperfunctional (PVH and NPVH) and normal vocal function. The findings provide direct validation of the use of ACC-based estimates of glottal aerodynamic measures to assess the pathophysiology of disordered voices and opens up the possibility of acquiring such measures in an unobtrusive way during natural connected speech, including during activities of daily living, which greatly expands the clinical and research capabilities of voice ambulatory monitors that use an ACC as the phonation sensor. Expanded capabilities for ambulatory monitoring of vocal function has the potential to provide better insight into the pathophysiological mechanisms associated with voice disorders, particularly for disorders such as those associated with VH in which daily voice use is assumed to play a role. Future efforts may further validate the use of ACC-based estimates of glottal aerodynamic measures to differentiate normal and pathological vocal function during natural connected speech.

## Author Contributions

**Victor M. Espinoza:** Conceptualization (Lead), Data curation (Lead), Formal analysis (Lead), Software (Lead), Writing - Original Draft (Lead), Writing—Review & Editing (Lead). **Daryush D. Mehta:** Conceptualization (Lead), Resources (Equal), Writing—Review & Editing (Lead). **Jarrad H. Van Stan:** Formal analysis (Supporting). **Robert E. Hillman:** Conceptualization (Lead), Project administration (Lead), Writing—Review & Editing (Lead). **Matías Zañartu:** Conceptualization (Lead), Project administration (Lead), Writing—Original Draft (Supporting), Writing—Review & Editing (Supporting).

## Acknowledgments

This research was supported in part by grants from the Voice Health Institute, the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders (Grants R33 DC011588, P50 DC015446 [awarded to Robert E. Hillman], and R21 DC015877 [awarded to Daryush D. Mehta]), ANID (Grants FONDECYT 1191369 and BASAL FB0008 [awarded to Matías Zañartu]), UTFSM (FSM1204 [awarded to Victor M. Espinoza]). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

- Bhattacharyya, N.** (2014). The prevalence of voice problems among adults in the United States. *The Laryngoscope*, *124*(10), 2359–2362. <https://doi.org/10.1002/lary.24740>
- Castellana, A., Carullo, A., Corbellini, S., & Astolfi, A.** (2018). Discriminating pathological voice from healthy voice using cepstral peak prominence smoothed distribution in sustained vowel. *IEEE Transactions on Instrumentation and Measurement*, *67*(3), 646–654. <https://doi.org/10.1109/TIM.2017.2781958>
- Cohen, J.** (1988). *Statistical power analysis for the behavior science* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Cortés, J. P., Espinoza, V. M., Ghassemi, M., Mehta, D. D., Van Stan, J. H., Hillman, R. E., Guttag, J. V., & Zañartu, M.** (2018). Ambulatory assessment of phonotraumatic vocal hyperfunction using glottal airflow measures estimated from neck-surface acceleration. *PLOS ONE*, *13*(12), 1–22. <https://doi.org/10.1371/journal.pone.0209017>
- Espinoza, V. M., Zañartu, M., Van Stan, J. H., Mehta, D. D., & Hillman, R. E.** (2017). Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2159–2169. [https://doi.org/10.1044/2017\\_JSLHR-S-16-0337](https://doi.org/10.1044/2017_JSLHR-S-16-0337)
- Fryd, A. S., Van Stan, J. H., Hillman, R. E., & Mehta, D. D.** (2016). Estimating subglottal pressure from neck-surface acceleration during normal voice production. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1335–1345. [https://doi.org/10.1044/2016\\_JSLHR-S-15-0430](https://doi.org/10.1044/2016_JSLHR-S-15-0430)
- Galindo, G. E., Peterson, S. D., Erath, B. D., Castro, C., Hillman, R. E., & Zañartu, M.** (2017). Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds. *Journal of Speech, Language, and Hearing Research*, *60*(9), 2452–2471. [https://doi.org/10.1044/2017\\_JSLHR-S-16-0412](https://doi.org/10.1044/2017_JSLHR-S-16-0412)
- Ghassemi, M., Van Stan, J. H., Mehta, D. D., Zañartu, M., Cheyne, H. A., II, Hillman, R. E., & Guttag, J. V.** (2014). Learning to detect vocal hyperfunction from ambulatory neck-surface acceleration features: Initial results for vocal fold nodules. *IEEE Transactions on Biomedical Engineering*, *61*(6), 1668–1675. <https://doi.org/10.1109/TBME.2013.2297372>
- 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*(2), 373–392. <https://doi.org/10.1044/jshr.3202.373>
- Hogikyan, N. D., & Sethuraman, G.** (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice*, *13*(4), 557–569. [https://doi.org/10.1016/S0892-1997\(99\)80010-1](https://doi.org/10.1016/S0892-1997(99)80010-1)
- Kempster, G. B., Gerratt, B. R., Abbott, K. V., 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*(2), 124–132. [https://doi.org/10.1044/1058-0360\(2008\)08-0017](https://doi.org/10.1044/1058-0360(2008)08-0017)
- Kennedy, J., & Eberhart, R.** (1995). Particle swarm optimization. *Proceedings of ICNN'95-International Conference on Neural Networks*, *4*, 1942–1948. <https://doi.org/10.1109/ICNN.1995.488968>
- Lei, Z., Kennedy, E., Fasanella, L., Li-Jessen, N. Y., & Mongeau, L.** (2019). Discrimination between modal, breathy and pressed voice for single vowels using neck-surface vibration signals. *Applied Sciences*, *9*(7), 1505. <https://doi.org/10.3390/app9071505>
- Lin, J. Z., Espinoza, V. M., Marks, K. L., Zañartu, M., & Mehta, D.** (2019). Improved subglottal pressure estimation from neck-surface



- vibration in healthy speakers producing non-modal phonation. *IEEE Journal of Selected Topics in Signal Processing*, 14(2), 449–460. <https://doi.org/10.1109/JSTSP.2019.2959267>
- Lico, A., Zañartu, M., González, A., Wodicka, G. R., Mehta, D. D., Van Stan, J. H., & Hillman, R. E.** (2015). Real-time estimation of aerodynamic features for ambulatory voice biofeedback. *JASA Express Letters*, 138(1), EL14–EL19. <https://doi.org/10.1121/1.4922364>
- Marks, K. L., Lin, J. Z., Fox, A. B., Toles, L. E., & Mehta, D. D.** (2019). Impact of nonmodal phonation on estimates of subglottal pressure from neck-surface acceleration in healthy speakers. *Journal of Speech, Language, and Hearing Research*, 62(9), 3339–3358. [https://doi.org/10.1044/2019\\_JSLHR-S-19-0067](https://doi.org/10.1044/2019_JSLHR-S-19-0067)
- Massey, F. J.** (1951). The Kolmogorov–Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253), 68–78. <https://doi.org/10.2307/2280095>
- McKenna, V. S., Lico, A. F., Mehta, D. D., Perkell, J. S., & Stepp, C. E.** (2017). Magnitude of neck-surface vibration as an estimate of subglottal pressure during modulations of vocal effort and intensity in healthy speakers. *Journal of Speech, Language, and Hearing Research*, 60(12), 3404–3416. [https://doi.org/10.1044/2017\\_JSLHR-S-17-0180](https://doi.org/10.1044/2017_JSLHR-S-17-0180)
- Mehta, D. D., & Hillman, R. E.** (2012). Current role of stroboscopy in laryngeal imaging. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 20(6), 429–436. <https://doi.org/10.1097/MOO.0b013e3283585f04>
- Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., Cortés, J. P., Cheyne, H. A., II., & Hillman, R. E.** (2015). Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in Bioengineering and Biotechnology*, 3(155). <https://doi.org/10.3389/fbioe.2015.00155>
- Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R.** (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology*, 27(3), 887–905. [https://doi.org/10.1044/2018\\_AJSLP-17-0009](https://doi.org/10.1044/2018_AJSLP-17-0009)
- 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(4), 1777–1781. <https://doi.org/10.1121/1.401011>
- 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(6), 1632–1645. <https://doi.org/10.1121/1.1913513>
- Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M.** (2005). Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *The Laryngoscope*, 115(11), 1988–1995. <https://doi.org/10.1097/01.mlg.0000179174.32345.41>
- Stevens, K. N.** (2000). *Acoustic phonetics*. MIT Press. <https://doi.org/10.7551/mitpress/1072.001.0001>
- Švec, J. G., Titze, I. R., & Popolo, P. S.** (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *The Journal of the Acoustical Society of America*, 117(3), 1386–1394. <https://doi.org/10.1121/1.1850074>
- Van Stan, J. H., Mehta, D. D., Ortiz, A. J., Burns, J. A., Toles, L. E., Marks, K. L., Vangel, M., Hron, T., Zeitels, S., & Hillman, R. E.** (2020). Differences in weeklong ambulatory vocal behavior between female patients with phonotraumatic lesions and matched controls. *Journal of Speech, Language, and Hearing Research*, 63(2), 372–384. [https://doi.org/10.1044/2019\\_JSLHR-19-00065](https://doi.org/10.1044/2019_JSLHR-19-00065)
- Van Stan, J. H., Mehta, D. D., Sternad, D., Petit, R., & Hillman, R. E.** (2017). Ambulatory voice biofeedback: Relative frequency and summary feedback effects on performance and retention of reduced vocal intensity in the daily lives of participants with normal voices. *Journal of Speech, Language, and Hearing Research*, 60(4), 853–864. [https://doi.org/10.1044/2016\\_JSLHR-S-16-0164](https://doi.org/10.1044/2016_JSLHR-S-16-0164)
- 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(6), 3262–3271. <https://doi.org/10.1121/1.4901714>
- 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(9), 1929–1939. <https://doi.org/10.1109/TASL.2013.2263138>