ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Estimating the Pathophysiology of Phonotraumatic Vocal Hyperfunction Using Ambulatory Data and a Computational Model

Jesús A. Parra,[a] Emiro J. Ibarra,[a] Carlos Calvache,[b] Jarrad H. Van Stan,[c] Robert E. Hillman,[c] and Matías Zañartu[a]

[a] Universidad Técnica Federico Santa María, Valparaíso, Chile [b] Corporación Universitaria Iberoamericana, Bogotá, Colombia [c] MGH Institute of Health Professions, Boston, MA

ARTICLE INFO

Article History:
Received June 11, 2024
Revision received September 15, 2024
Accepted November 19, 2024

Editor-in-Chief: Rita R. Patel
Editor: Susan L. Thibeault

https://doi.org/10.1044/2024_JSLHR-24-00419

ABSTRACT

Purpose: This study uses a voice production model to estimate muscle activation levels and subglottal pressure (P_S) in patients with phonotraumatic vocal hyperfunction (PVH), based on ambulatory measurements of sound pressure level (SPL) and spectral tilt (H_1–H_2). In addition, variations in these physiological parameters are evaluated with respect to different values of the Daily Phonotrauma Index (DPI).
Method: The study obtained ambulatory voice data from patients diagnosed with PVH and a matched control group. To infer physiological parameters, ambulatory data were mapped onto synthetic data generated by a physiologically relevant voice production model. Inverse mapping strategies involved selecting model simulations that represented ambulatory distributions using stochastic (random) sampling weighted by probability with which different vowels occur in English. A categorical approach assessed the relationship between different values of DPI and changes in estimated physiological parameters.
Results: Results showed significant differences between the PVH and control groups in key parameters, including statistical moments of H_1–H_2, SPL, P_S, and muscle activity of lateral cricoarytenoid (LCA) and cricothyroid (CT) muscles. Higher DPI values, reflecting more severe PVH, were associated with increased mean LCA activation and decreased LCA variability, along with decreased mean CT activation and increased median P_S. These findings highlight the relationship between muscle activation patterns, P_S, and the severity of vocal pathology as indicated by the DPI. It is hypothesized that a major driver of muscle activation and P_S changes is the variation in maladaptive adjustments (vocal effort) when compensating for the presence of vocal pathology.
Conclusions: This study demonstrated that noninvasive ambulatory voice data could be used to drive a voice production modeling process, providing valuable insights into underlying physiological parameters associated with PVH. Future research will focus on refining the predictive power of the modeling process and exploring the implications of these findings in further delineating the etiology and pathophysiology of PVH, with the ultimate goal to develop improved methods for the prevention, diagnosis, and treatment of PVH.
Supplemental Material: https://doi.org/10.23641/asha.28352720

Correspondence to Jesús A. Parra: jesus.parrap@sansano.usm.cl. ***Disclosure:*** *Carlos Calvache has financial interests in Vocology Center, a company focused on vocal rehabilitation processes, professional voice training, and continuing education in the field of spoken, sung, and interpreted voice. Calvache's interests were reviewed and are managed by Corporación Universitaria Iberoamericana, Bogotá, Colombia, in accordance with its conflict-of-interest policies. Robert E. Hillman has a financial interest in InnoVoyce LLC, a company focused on developing and commercializing technologies for the prevention, diagnosis, and treatment of voice-related disorders. Hillman's interests were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies. Matías Zañartu has a financial interest in Lanek SPA, a company focused on developing and commercializing biomedical devices and technologies. Zañartu's interests were reviewed and are managed by Universidad Técnica Federico Santa María in accordance with its conflict-of-interest policies. All other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Voice disorders affect approximately 25%–30% of the adult population in the United States (Bhattacharya & Siegmund, 2015; Roy et al., 2005), with some of the most common being associated with vocal behaviors that cause vocal fold trauma (e.g., nodules and polyps; Hillman et al., 1989, 2020)—referred to as phonotraumatic vocal hyperfunction (PVH). Phonation-related vocal fold trauma can degrade vocal function/efficiency (e.g., interfere with glottal closure) and require increased vocal effort during daily vocal use to compensate (Behrman et al., 2008; Dikkers & Nikkels, 1995; Perkell et al., 1994). Such circumstances have the potential to trigger a vicious cycle of increasing trauma and effort (Hillman et al., 2020).

Because PVH is related to daily vocal behaviors, ambulatory monitoring is considered an ideal tool for characterizing voice use in such patients (Carullo et al., 2013; Cheyne et al., 2003; Popolo et al., 2005; Searl & Dietsch, 2014). Monitoring devices typically employ sensors placed on the neck, such as miniaturized accelerometers (ACCs), to detect neck skin vibration and, thus, noninvasively monitor phonation (Cheyne et al., 2003; Popolo et al., 2005).

A recently developed measure called the Daily Phonotrauma Index (DPI) differentiated the daily voice use of patients with PVH from matched controls (Nudelman et al., 2022; Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020; Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020). The DPI is a logistic regression that combines sound pressure level (SPL; Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020) or neck skin acceleration magnitude (NSAM; Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020) with the amplitude differences between the first and second harmonics of the voice spectrum (H_1–H_2; Cortés et al., 2018; Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020). Patients with PVH have more negatively skewed SPL/NSAM distributions (hypothesized to reflect increased laryngeal forces) and less variable H_1–H_2 distributions centered toward lower values (hypothesized to reflect more abrupt glottal closure).

Although the DPI has been shown to differentiate between patients with PVH and controls (Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020) and correlate with treatment-related changes in PVH (Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020; Van Stan, Ortiz, Marks, et al., 2021), differences in physiological mechanisms, such as laryngeal muscle activity/tension or subglottal pressure (P_S), that underlie typical versus aberrant DPI values have not been identified/validated. Identifying/validating these mechanisms would greatly increase the interpretability and clinical value of the DPI. Because the DPI is based on measures that are extracted from noninvasive ambulatory voice data, it would be ideal if a voice production modeling process could be developed that uses these measures as a basis for estimating underlying physiological parameters.

Several studies have shown promise in the noninvasive estimation of P_S using various methodological strategies. These include nonlinear regressions trained with numerical voice production models (Gomez et al., 2019; Ibarra et al., 2021; Zhang, 2020, 2022), empirical formulas correlating P_S with SPL and fundamental frequency ($F0$; Titze & Hunter, 2015; Titze & Sundberg, 1992), or exploiting linear correlations between root-mean-square of the ACC signal and P_S (Fryd et al., 2016; Lin et al., 2020; Marks et al., 2019, 2020). The use of a neck-placed ACC sensor in several of these efforts (Ibarra et al., 2021; Sepúlveda et al., 2024) has shown its value as an instrument for measuring vocal function. Recent work has used the ACC for ambulatory recordings, thus enabling estimates of P_S distributions for the entire day by employing subject-specific linear regressions based on laboratory calibrations of the ACC (Cortés, Lin, et al., 2022).

In this study, we use mathematical models of voice production to begin bridging the gap between noninvasive voice ambulatory measures that differentiate patients with PVH from controls and physiological parameters that underlie these differences. These models allow for the simulation and study of interactions between anatomical, biomechanical (Calvache et al., 2023), and acoustic variables involved in phonation (Döllinger et al., 2023; Erath et al., 2013). For this purpose, a fully interactive implementation of the triangular body-cover model (TBCM; Alzamendi et al., 2022), a reduced-order mass-spring model of the vocal folds with laryngeal muscle control, is utilized as a voice production model for simulating different phonatory gestures that can be related to typical and hyperfunctional voice use (Alzamendi et al., 2022; Galindo et al., 2017; Zañartu et al., 2014). In addition, this model has successfully estimated laryngeal muscle activation levels and P_S, contrasted with clinical data (Alzamendi et al., 2020; Ibarra et al., 2021), thus making it an ideal model for the purpose of this study.

This work has two aims. First, we will replicate patient and healthy control ambulatory distributions underlying the DPI (H_1–H_2 and SPL) with a TBCM to estimate differences in muscle activation levels and P_S. It is hypothesized that P_S, muscle activation levels, and adduction/glottal closure will be higher in the patients than controls. Second, we will replicate five ambulatory distributions along the severity continuum of the DPI (as a probability). We expect the differences in muscle activation and P_S, between patients and controls from the first aim, to increase or decrease along with the DPI. To validate these hypotheses, we will statistically compare the resulting distributions and features between controls and patients with paired t tests and Cohen's d values (Cohen, 1988). This study represents a synergistic approach, integrating ambulatory data from the PVH population, mathematical models of phonation (using the TBCM), and

statistical analysis tools (using DPI) to identify the underlying elements of vocal hyperfunction in daily life.

Method

The proposed methodology for linking ambulatory data distributions with the voice production model is illustrated in Figure 1. This figure encapsulates the convergence of the three pivotal elements required to generate P_S and muscle activation distributions based on SPL and H_1–H_2 distributions. As shown, the three components include the ambulatory data, the simulations from the model, and the set of guidelines devised to mimic different vowels in the ambulatory environment using the model, that is, the inverse mapping scheme.

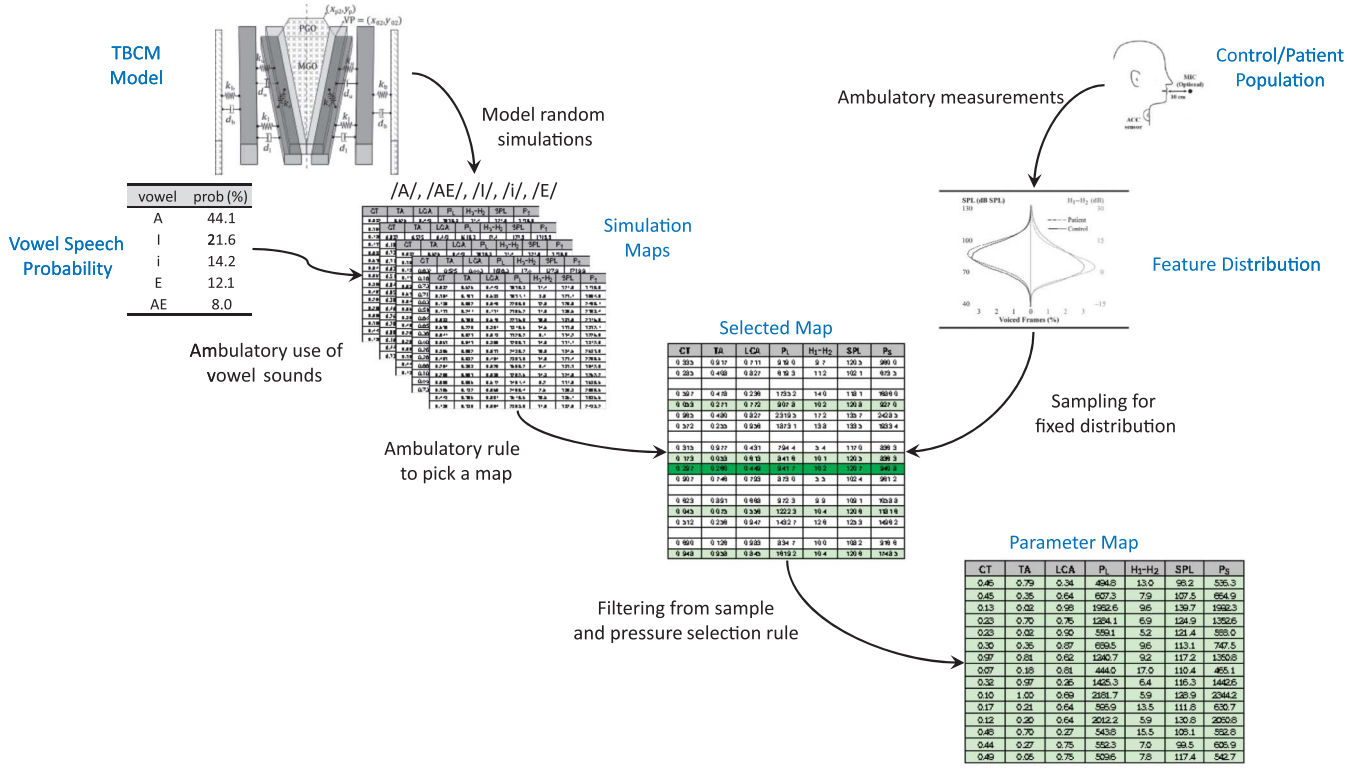
Ambulatory Distributions of Subjects With PVH and Controls

The acquisition and composition of the ambulatory database employed in this study have been previously described (Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020). This data set included recordings from 180 adult women—90 of whom were patients diagnosed with PVH (vocal fold nodules or polyps) and another 90 who acted as control subjects matched to the patients in terms of age, sex, and occupation and endoscopically screened to ensure the absence of vocal pathology. The governing institutional review board (IRB) from the Massachusetts General Hospital approved all experimental aspects related to the use of human subjects for this study in IRB Protocol 2011P002376.

Each participant was recorded for an entire week using the Voice Health Monitor (VHM), as described in Mehta et al. (2012). To briefly summarize, the VHM uses a miniature/lightweight ACC (Model BU-27135, Knowles Electronics), placed on the anterior neck—just below the larynx and above the sternal notch—to sense the skin surface vibrations that are generated during phonation. The ACC is connected to a smartphone that runs an app for data acquisition and storage. In addition, a handheld microphone (H1 Handy Recorder, Zoom Corporation) was utilized for calibration of SPL once a day.

From the ambulatory voice data for each participant, we derived distributions and summary statistics for

Figure 1. Diagram of the inverse mapping scheme. The right track delineates ambulatory data and distributions from both subjects with phonotraumatic vocal hyperfunction and controls. The left track showcases data from triangular body-cover model (TBCM) simulations. At the center, the figure emphasizes the filtering of simulated data, grounded in established distributions. MGO = membranous glottal opening; PGO = posterior glottal opening; ACC = accelerometer; CT = cricothyroid; LCA = lateral cricoarytenoid; P_L = lung pressure; prob = probability; SPL = sound pressure level; TA = thyroarytenoid.



the entire week for 50-ms frames of SPL, $F0$, and H_1–H_2. The summary statistics included the mean, standard deviation, skewness, and kurtosis. To convert the ACC signal into a glottal flow signal and subsequently quantify the aerodynamic features, we implemented the Impedance-Based Inverse Filtering (IBIF) algorithm as described by Zañartu et al. (2013) for each subject. A subject-specific linear regression was used to transform NSAM into estimated SPL values (Mehta et al., 2012).

Phonation Model

The TBCM (Alzamendi et al., 2022) is a reduced-order biomechanical model that simulates symmetrical oscillatory behavior of the vocal folds during phonation. The model distinguishes between the body (represented by a single mass) and the cover layers (represented by two masses) of the vocal folds (Story & Titze, 1995) and accurately reflects the triangular glottal configuration observed in natural phonation (Galindo et al., 2017). It incorporates the coordinated activation of the five intrinsic laryngeal muscles—thyroarytenoid (TA), cricothyroid (CT), lateral cricoarytenoid (LCA), interarytenoid, and posterior cricoarytenoid (PCA)—each influencing vocal fold stiffness, mass distribution, and posture, thereby impacting pitch, loudness, and phonatory quality (Titze & Hunter, 2007; Titze & Story, 2002). These five intrinsic laryngeal muscles have normalized values between 0 (rest) and 1 (tetanus, full contraction).

The ambulatory ACC data are known to be influenced by the subglottal and supraglottal systems (Zañartu et al., 2013). To account for these complex interactions between tissue, airflow, and acoustic waves above and below the vocal folds, we integrated a vocal tract model that simulates the propagation of acoustic waves in the time domain through both the subglottal and supraglottal systems using a wave reflection analog scheme (Zañartu et al., 2014). A schematic representation of the TBCM is shown in the top left corner of Figure 1.

The model simulates sustained phonation segments of 0.5-s duration with only one vowel (i.e., supraglottal vocal tract) and vocal register (i.e., TBCM) per trial. However, the ambulatory data consist of spontaneous speech in daily life (i.e., many different vowels) and multiple vocal registers. Thus, to achieve different vowels and multiple modal registers (variations in pitch, loudness, and phonatory quality), the model uses input time-varying parameters such as the supraglottal tract cross-sectional area for a given vowel, normalized muscle activation levels (a_{CT}, a_{TA}, a_{LCA}, a_{IA}, and a_{PCA}), and lung pressure (P_L). A truncated Taylor series approximation is implemented to simultaneously solve the differential equations of motion for the three masses that compose the body-cover structure, using a sampling frequency of 44.1 kHz. The vocal tract area functions are also set according to Story (2008) to mimic several vowels.

For each simulation, the output signals of interest include the P_S, oral airflow, and radiated sound pressure. The $F0$ and H_1–H_2 features are computed from the glottal airflow signal, the P_S is derived as the mean of P_S signal, and the SPL is obtained from the radiated pressure. Notably, the $F0$, H_1–H_2, and SPL features are synonymous with those derived from the individual ACC data and the IBIF algorithm.

Synthetic Data Set

To make the model output tractable, we focused on the five most commonly used vowel sounds in American English: /I/, /i/, /e/, /AE/, and /A/. We performed 500,000 simulations for each vocal tract configuration. These simulations varied the remaining input model parameters (a_{CT}, a_{TA}, a_{LCA}, and P_L) using a random parametric sweep within the ranges specified in Table 1. This approach ensured a dense and representative exploration of the model's input space. Only simulations that produced sustained oscillatory behavior were retained, resulting in a total synthetic data set comprising 1.5 million simulations. Each individual simulation was conducted with constant values for muscle activation and P_L. After 1 s of simulation, the last 100 ms were used to compute the $F0$, H_1–H_2, P_S, and the SPL. To streamline the parameter space, the activation level of the PCA muscle (a_{PCA}) was consistently set to 0.

We constructed five vowel maps, each representing a distinct vocal tract configuration. These maps consist of comprehensive sets of unique input vectors (a_{CT}, a_{TA}, a_{LCA}, and P_L) and their corresponding output vectors (SPL, H_1–H_2, P_S) for each specified vocal tract shape, as illustrated in Figure 1.

Inverse Mapping Procedure

The integration of experimental and simulated data forms the cornerstone for deriving parameter distributions

Table 1. Input parameters for triangular body-cover model simulations.

Input parameter	Acronym	Range
Lung pressure	P_L	500–2,500 Pa
Act. of cricothyroid	a_{CT}	0–1
Act. of thyroarytenoid	a_{TA}	0–1
Act. of lateral cricoarytenoid	a_{LCA}	0–1
Act. of interarytenoid	a_{IA}	$a_{IA} = a_{LCA}$
Act. of posterior cricoarytenoid	a_{PCA}	0
Vocal tract	VT	/I/, /i/, /e/, /AE/, /A/

Note. Act = activation.

through an inverse mapping strategy, also known as the sample-and-filter method. This entails sampling the SPL and H_1–H_2 distributions from ambulatory data and subsequently filtering the simulation outputs to isolate input–output vectors corresponding to the sampled values.

To perform the sampling and filtering process, four steps are carried out. First, we pick a point from the SPL and H_1–H_2 distributions extracted from a participant's ambulatory data. Second, we use a sampling rule that assigns a probability to selecting each vowel, which is determined by the relative occurrence of each vowel in natural spoken English (Lammert et al., 2020; Parra et al., 2023). The probability of selection for each vowel was as follows: 44% for /A/, 21% for /I/, 14% for /i/, 12% for /E/ and 8% for /AE/. Third, from the selected vowel map, we identify the vectors that have SPL and H_1–H_2 values that approximate the points taken from the ambulatory data distributions. The tolerance levels for identifying approximate vectors were 0.3 dB for SPL and 0.2 dB for H_1–H_2 when calculating the distance to the chosen point. These tolerance levels were defined based on the 1% width of the distribution of values of each feature. Fourth, within the subset of filtered results, the final input–output pair that reproduces the sampled point is randomly selected. We iterated this process 20,000 times to acquire a collection of vectors, which not only aligned with the subject-specific feature distributions (SPL, H_1–H_2) but also revealed the model parameter values responsible for producing those features. If there was an individual veridical data point that was not approximated by any SPL–H_1-H_2 combinations, these data points were not included in the final simulated distribution.

Statistical Analysis

Similar to the approach used in previous work (Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020; Van Stan, Ortiz, Cortes, et al., 2021), we compared distributions of SPL, H_1–H_2, muscle activations, and P_S between patients and controls (Aim 1). The analysis included central tendency and dispersion metrics—mean, median, standard deviation—as well as measures of distribution shape, namely, skewness and kurtosis, for each feature within each subject data set (SPL, H_1–H_2, a_{CT}, a_{TA}, a_{LCA}, and P_S), to a total of 30 features per subject.

To analyze the matched control–patient paradigm (90 pairs), paired t tests (parametric data) and Wilcoxon signed-ranks tests (nonparametric data) were used to evaluate differences in 30 resulting features. A Kolmogorov–Smirnov (KS) test assessed the normalcy of each distribution of paired differences (patient minus control). If the KS test was significant ($p < .05$), a Wilcoxon test was applied; if not, a paired t test was used. Due to the high number of tests, the alpha significance level was adjusted using the Bonferroni method. Upon finding statistical significance, the Cohen's d effect size was calculated to characterize the magnitude of differences between groups (small when ≤ 0.19, small to medium when 0.20–0.49, medium to large when 0.50–0.79, and large when ≥ 0.80; Cohen, 1988).

For Aim 2, we examined parameters that were significantly different between the control–patient pairings across five DPI severity ranges. A DPI value was determined for each participant using H_1–H_2 standard deviation and SPL skewness in the calculation first described by Van Stan, Mehta, Ortiz, Burns, Marks, et al. (2020). For the analysis, five DPI groups were defined, each representing a probability range of .2 within the overall DPI scale ($0 < DPI < 1$). This ensured that each group had the same probability range and a similar number of subjects (approximately 36 ± 2 per group). The average distributions for each parameter were plotted across the incremental levels of DPI using box plots, thus allowing the visualization of any relationships between a parameter and different DPI values.

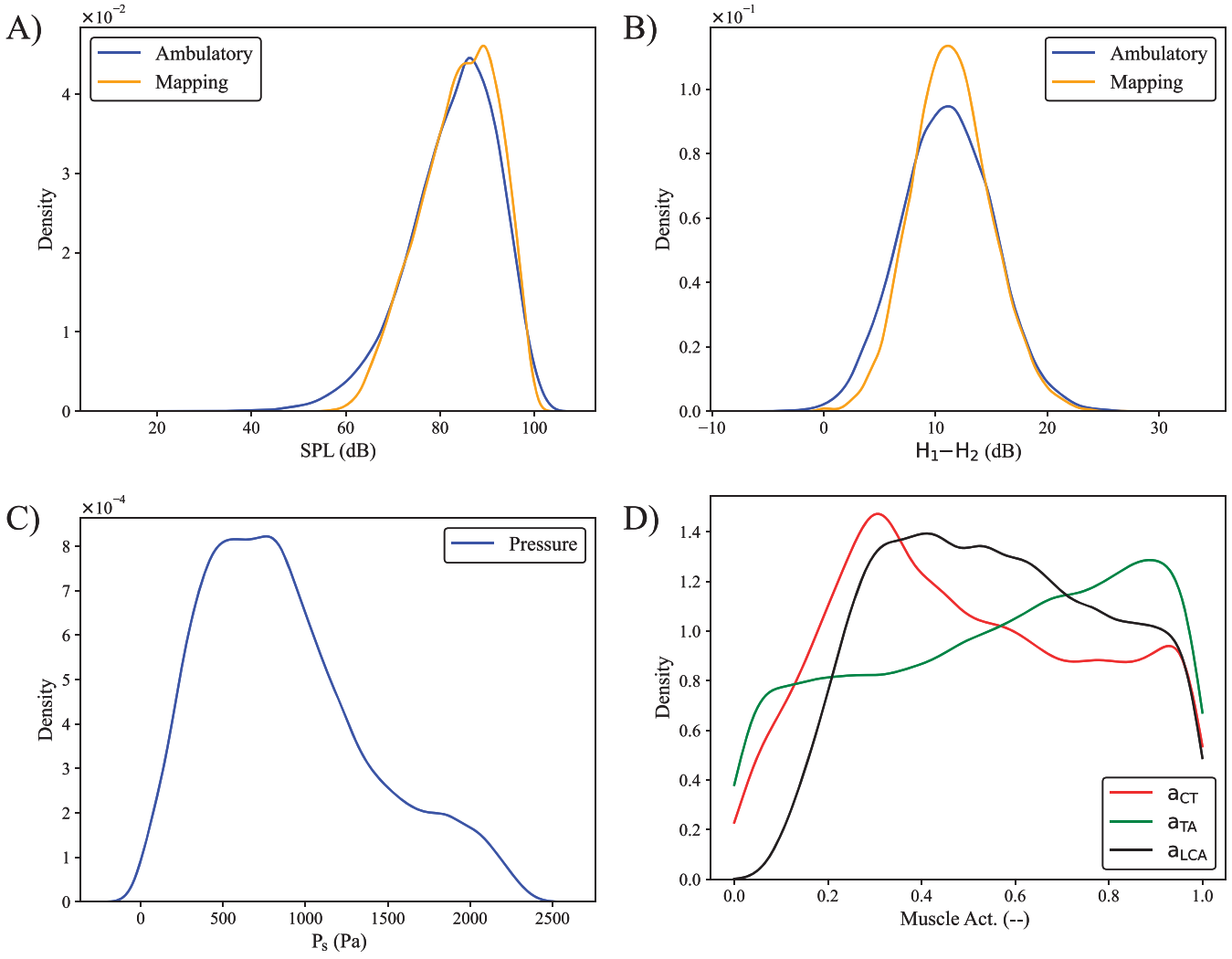
Results

Inverse Mapping Procedure

The inverse mapping procedure was successfully applied for all study subjects, resulting in model-produced SPL and H_1–H_2 distributions that closely matched the distributions based on the ambulatory SPL and H_1–H_2 data. Success in this context was defined as matching a significant proportion of the data points, with a threshold of approximately 60% considered reasonable for successful alignment. Notably, the actual success rate of the inverse mapping, measured by the proportion of matched points between ambulatory and model-derived SPL and H_1–H_2 pairs, was 77.8% ($\pm$6.1%), exceeding our threshold for success. This achievement is particularly noteworthy considering that the numerical model approximates the complex phonatory process, yet it replicated data across multiple individuals over three quarters of the time on average. This high success rate highlights a substantial congruence between the model outputs and the real-world voice data regarding these key voice features.

Figure 2 shows an example with feature distributions for one selected subject, both ambulatory and model derived. The upper panels display the SPL and H_1–H_2 distributions from both ambulatory data and the inverse mapping, illustrating a robust alignment between the observed and modeled data. The lower panels delineate the distributions of P_S and muscle activation inferred from the model. The mode of P_S, shown in Figure 2(C), is

Figure 2. Example of a subject-specific distribution of model parameters and features, after applying inverse mapping. Acoustic features: (A) sound pressure level (SPL): ambulatory (blue) and inverse mapping (orange) distribution. (B) H_1–H_2: ambulatory (blue) and inverse mapping (orange) distribution. Muscle activation (Act.) model parameters: (C) subglottal pressure (P_S) distribution. (D) Muscle activation: a_{TA} (red), a_{CT} (green), and a_{LCA} (gray) distribution. a_{CT} = activation of the cricothyroid muscle; a_{TA} = activation of the thyroarytenoid muscle; a_{LCA} = activation of the lateral cricoarytenoid muscle.



observed at approximately 0.7 kPa, with a mean around 0.9 kPa, corroborating with previously reported "comfortable loudness" measurements in controlled environments (Espinoza et al., 2017, 2020).

Regarding muscle activations, the distribution of the LCA muscle in Figure 2(D) shows a low or nearly zero occurrence of activity below 0.2. This is consistent with the essential role of LCA activation in producing phonation, given its adductor function (Chhetri et al., 2014). Conversely, the CT muscle distribution exhibits a higher prevalence of activation around 0.3, a range typically associated with a "speech" type of voice (Vahabzadeh-Hagh et al., 2018). The CT muscle also maintains activation across its entire range, reflecting the variability in vocal frequency. Finally, the TA muscle distribution shows activity throughout its range, with a preference for high activation levels. This trend should be interpreted in light of the inverse mapping and filtering scheme applied in this study. Previous experiments and regression analyses (Chhetri et al., 2014; Ibarra et al., 2021) have not demonstrated a strong correlation between TA activation and commonly used acoustic or aerodynamic parameters, resulting in a high degree of uncertainty in its estimation. Additionally, high TA activation is known to produce midfrequency phonation, while low activation can inhibit vocal fold oscillation in certain configurations (Chhetri et al., 2014). These dynamics influence the inverse mapping process, leading to a tendency to select more simulations with high TA activation. These observed trends are

consistent across other subjects. Moving forward, we transition to a more population-based analysis to further explore these findings.

Statistical Comparison of the PVH and Control Groups

As outlined in the methodology, measures of central tendency and statistical moments (mean, median, standard deviation, skewness, and kurtosis) were calculated for each subject's distributions (SPL, H_1–H_2, a_{CT}, a_{TA}, a_{LCA}, and P_S), resulting in a total of 30 parameters. Table 2 lists the parameters' means and standard deviations that statistical testing (KS test, Wilcoxon sum statistic, and Cohen's *d*) showed were significantly different between the PVH and control groups. Significant differences were found for the two parameters that comprise the DPI (H_1–H_2 standard deviation and SLP skewness), several distributional parameters for LCA muscle activation (mean, median, standard deviation), the mean for CT muscle activation, and the median for P_S. Not surprisingly, the two DPI components had the largest values of Cohen's *d*, with H_1–H_2 standard deviation having a large effect size (d = 1.12) and SLP skewness having a medium-to-large effect size (d = 0.63). The distributional features of LCA muscle activation showed medium effect sizes (|d| = 0.53–0.58), and both CT muscle activation (|d| = 0.58–0.43) and P_S had small-to-medium (d = 0.4) effective sizes. There were no significant parameters for TA muscle activation. Conversely, the significant parameters associated with the CT and LCA muscles underscore their correlation with elements of harmonic richness (H_1–H_2). The LCA muscle, as the primary adductor muscle, is crucial for vocal fold closure and posture, whereas the CT muscle, as the vocal fold tensor, plays a key role in pitch coordination. Additionally, the close relationship between P_S and SPL is further confirmed. The table containing each subject's central tendency and statistical moment values is provided in Supplemental Material S1 for controls and Supplemental Material S2 for patients.

Relationships Between Significant Parameters and DPI

Figures 3–5 display box plots of the statistically significant parameters (see Table 2) shown across increasing values for DPI. Higher DPI values are generally viewed as reflecting more severe PVH. The solid (99%) and dashed (95%) lines in the head of the box plots represent statistical difference between groups. All of the displayed parameters showed significant changes when comparing their distributions at a DPI value of 0.2 with those at both 0.8 and 1.0, demonstrating the ability to discriminate between higher control–patient contrasts.

Not surprisingly, the box plots in Figure 3 for H_1–H_2 standard deviation and SPL skewness display significant changes in these two components of the DPI across the entire range of DPI values, with the decrease in H_1–H_2 standard deviation being even more consistent than the increase in negative skewness of SPL. These results serve as a simple sanity check.

Box plots for the model-derived mean and standard deviation of LCA muscle activation are shown in Figure 4. In general, the mean LCA activation level increases significantly and the standard deviation of the LCA activation level decreases significantly with increases in DPI. This can be interpreted as more vocal fold closure on average and reduced variation toward less vocal fold closure (respectively) as the DPI increases.

Figure 5 displays the model-derived values for mean CT muscle activation and the median P_S. Overall, there is a significant but inconsistent decrease in the mean CT muscle activation level and a significant but inconsistent increase in median P_S with increases in DPI, indicating higher aerodynamic forces when the DPI increases.

Figure 6 illustrates the average distribution for both ambulatory and model-based parameters across the two DPI groups: DPI > 0.5 (classified as patients) and DPI <

Table 2. Features that showed significant differences between controls and patients with phonotraumatic vocal hyperfunction, with their *p* values and effect sizes.

Feature	Control (*M* ± *SD*)	Patient (*M* ± *SD*)	KS test (*p* value)	Wilcoxon (*p* value)	Cohen's *d*
H_1–H_2 *SD*	3.77 ± 0.33 dB	3.37 ± 0.38 dB	< .01	< .01	1.12
SPL skew	0.01 ± 0.24	–0.15 ± 0.22	< .01	< .01	0.63
a_{LCA} mean	0.57 ± 0.01	0.58 ± 0.02	< .01	< .01	–0.58
a_{LCA} median	0.56 ± 0.02	0.58 ± 0.02	< .01	< .01	–0.54
a_{LCA} *SD*	0.24 ± 0.01	0.23 ± 0.01	< .01	< .01	0.53
a_{CT} mean	0.50 ± 0.02	0.49 ± 0.02	< .01	< .01	0.43
P_S median	1.11 ± 0.15 kPa	1.17 ± 0.16 kPa	.02	< .01	–0.40

Note. KS = Kolmogorov–Smirnov; H_1–H_2 = difference between the first and second harmonics; SPL = sound pressure level; skew = skewness; a_{LCA} = activation of the lateral cricoarytenoid muscle; a_{CT} = activation of the cricothyroid muscle; P_S = subglottal pressure.

Figure 3. (Left) Standard deviation of H_1–H_2. (Right) Skewness of SPL. Solid (99%) and dashed (95%) lines represent statistical difference between groups. DPI = Daily Phonotrauma Index; *SD* = standard deviation; skew = skewness; SPL = sound pressure level; H_1–H_2 = difference between the first and second harmonics.

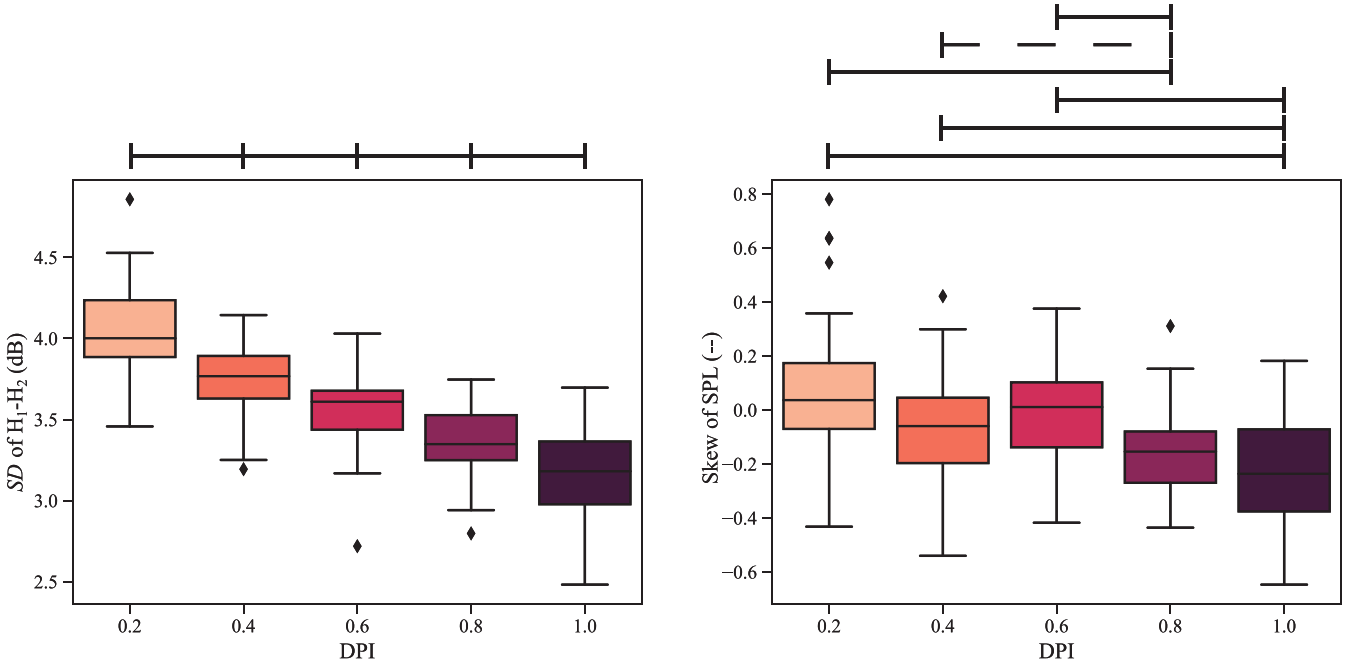


Figure 4. (Left) Mean of a_{LCA}. (Right) Standard deviation of a_{LCA}. Solid (99%) and dashed lines (95%) represent statistical difference between groups. DPI = Daily Phonotrauma Index; *SD* = standard deviation; a_{LCA} = activation of the lateral cricoarytenoid muscle.

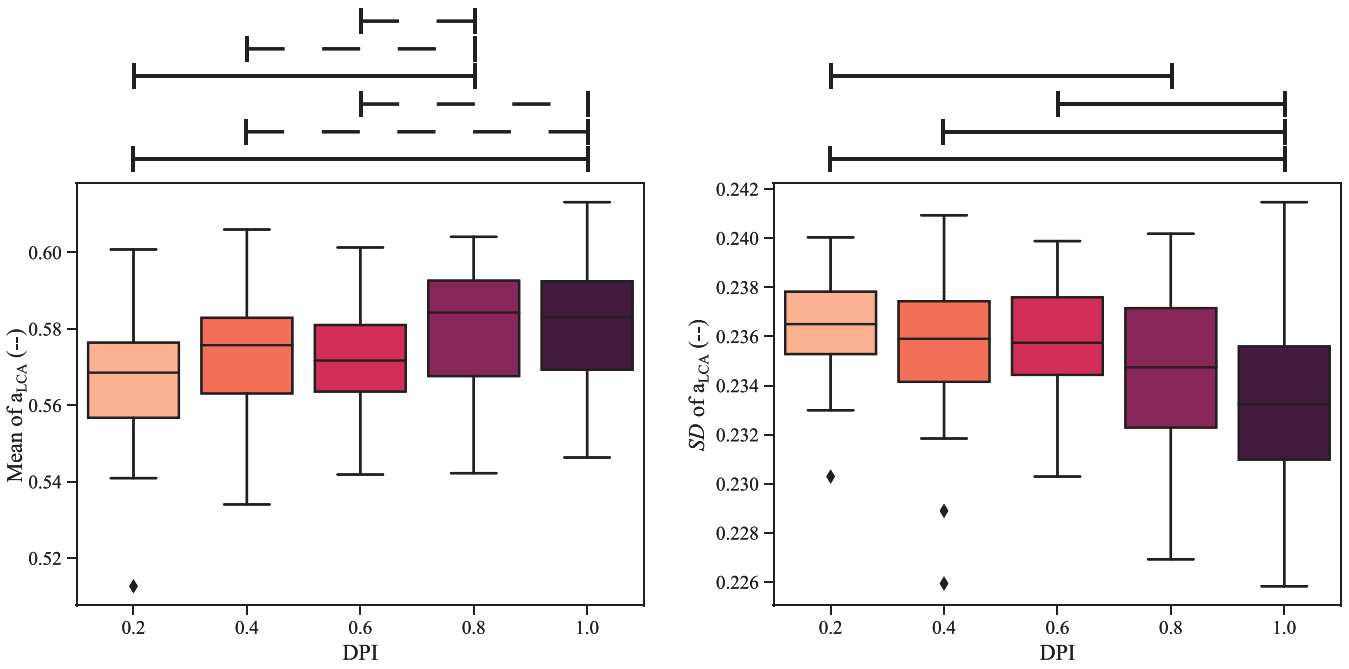


Figure 5. (Left) Mean of a_{CT}. (Right) Median of P_S. Solid (99%) and dashed (95%) lines represent statistical difference between groups. DPI = Daily Phonotrauma Index; a_{CT} = activation of the cricothyroid muscle; P_S = subglottal pressure.

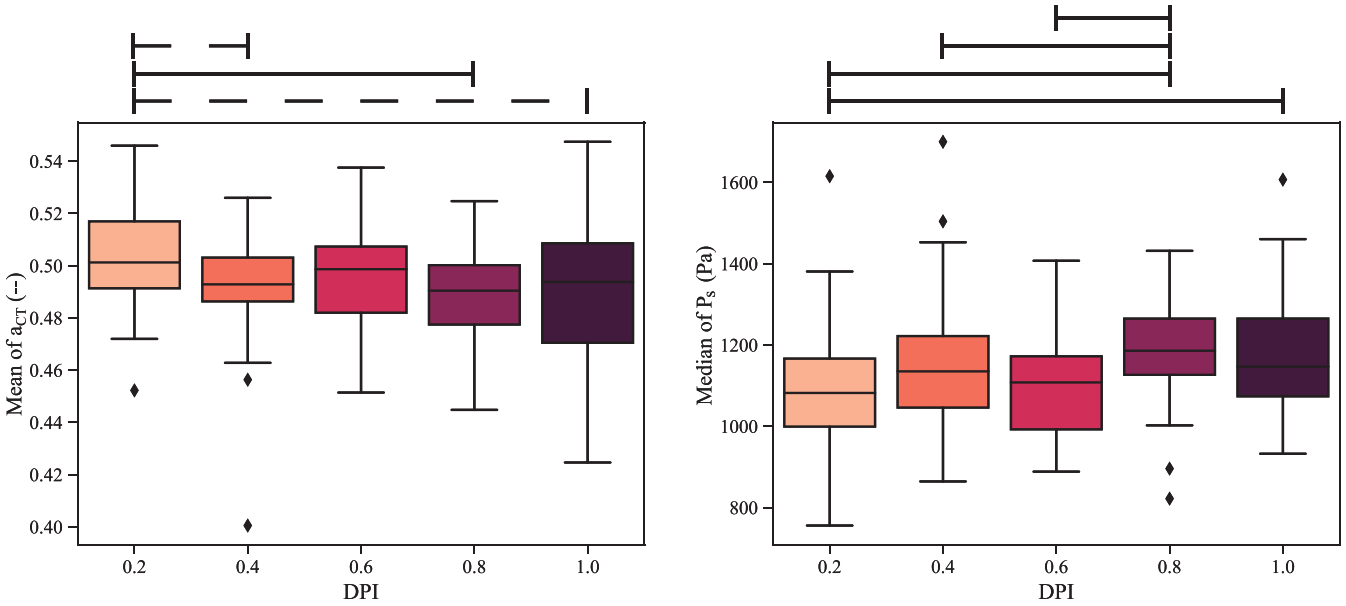
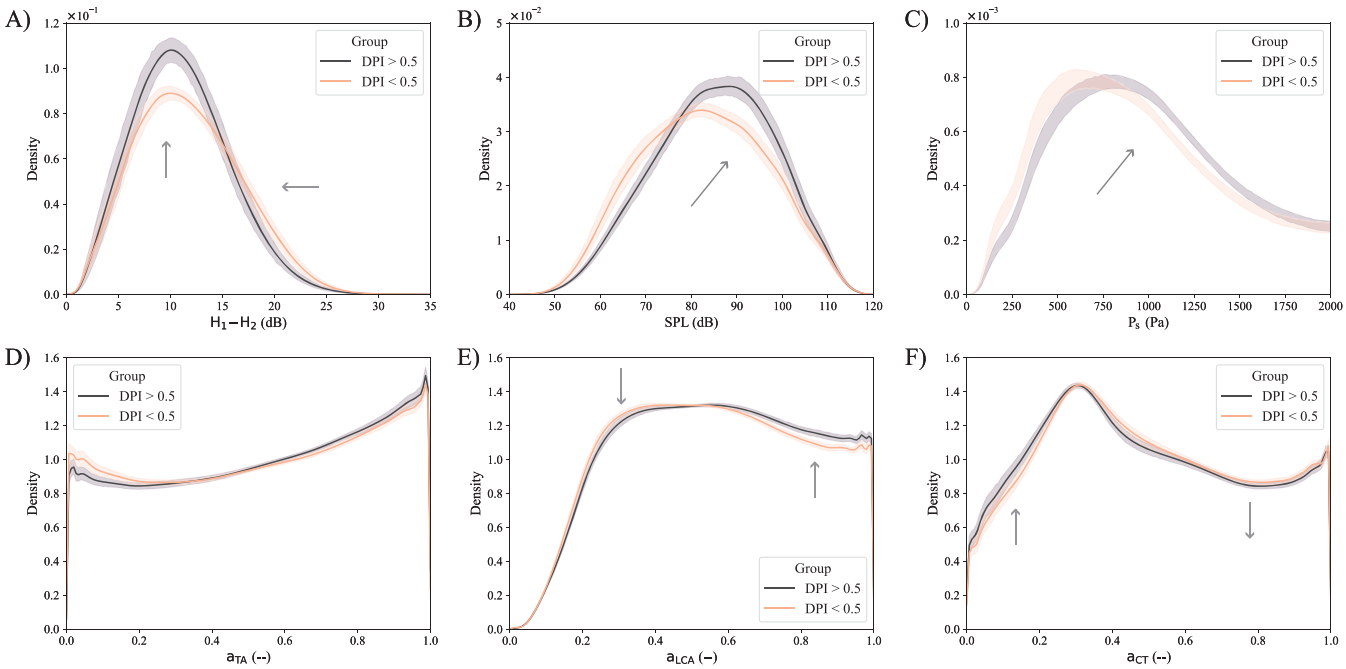


Figure 6. Mean distributions of parameters of interest. Acoustic features: (A) SPL and (B) H_1–H_2 distribution. Model parameters: (C) a_{TA}, (D) a_{LCA}, (E) a_{CT}, and (F) P_S distribution. Colors represent clustering by DPI value (groups of 0.2). The shaded area represents the 95% confidence interval. Gray arrows indicate the direction of DPI increase. DPI = Daily Phonotrauma Index; SPL = sound pressure level; H_1–H_2 = difference between the first and second harmonics; a_{CT} = activation of the cricothyroid muscle; a_{TA} = activation of the thyroarytenoid muscle; a_{LCA} = activation of the lateral cricoarytenoid muscle; P_S = subglottal pressure.



0.5 (classified as controls), summarizing the results from previous figures. The shaded areas represent the 95% confidence interval, and the gray arrows highlight the direction of increasing DPI scores.

For ambulatory parameters, the H_1–H_2 distribution, shown in Figure 6(A), becomes more concentrated, maintaining its tails while the central density heightens, thus diminishing the standard deviation when the DPI value increases. Concurrently, SPL distribution, shown in Figure 6(B), shifts toward elevated SPL values, indicating a skewness alteration. These elements represent what is observed in Figure 3 in terms of global parameters and not just their statistical moments. For model-derived parameters, P_S follows a similar trend to SPL, reflecting their correlation, whereas LCA activation increases with higher DPI scores, as seen in Figure 4. CT activation decreases as DPI increases, and TA activation shows no significant variation across DPI categories.

Discussion

Our study had two hypotheses: (a) P_S and muscle activation levels related to adduction/glottal closure will be higher in the patients than controls, and (b) the differences in muscle activation and P_S between patients and controls, from the first hypothesis, will increase or decrease along with the DPI. To evaluate any muscle activation or P_S values, the TBCM with subglottal and supraglottal tracts needed to replicate distributions of SPL and H_1–H_2. Notably, there was high agreement between the vocal production model predictions and actual ambulatory measurements of SPL and H_1–H_2. The strong consistency between ambulatory data and model predictions lends some validation to using the TBCM and associated lookup tables for estimating physiological parameters from the neck skin acceleration signal. To match real-life SPL and H_1–H_2 distributions, the TBCM and lookup tables must accurately reflect the physical processes of vocal production to a nontrivial degree. Regarding muscle activation, the distributions reflect a prevalence of lower activations (below 0.5) for the CT muscle and a predominance of higher activations (above 0.4) for the TA and LCA muscles, replicating the modal phonation pattern described in our previous TBCM article (Alzamendi et al., 2022).

The largest differences in estimated physiological measures were patients showing increased average LCA values (d = 0.54–0.58) and decreased LCA variability (d = 0.53) compared to controls. Excess average LCA activation would adduct the vocal folds more closely during voicing, increasing the vocal folds' tissue-to-tissue contact. It is thought that normal voicing occurs with the vocal folds barely touching (Verdolini et al., 1998). Patients with PVH increase their risk of phonotrauma by overadducting their vocal folds and/or voicing in ways that increase tissue-to-tissue contact (Erath et al., 2017; Mehta et al., 2021). Additionally, these patients have lesions in the midmembranous vocal fold, which would prevent full vocal fold closure during voicing and further necessitate increased LCA muscle activation/vocal fold adduction to produce clear voicing (Hillman et al., 2020). Reduced LCA variability corresponds to previous work with the DPI, showing that patients voiced with decreased variation to higher values of H_1–H_2 in daily life (Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020; Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020; Van Stan, Ortiz, Marks, et al., 2021), reflecting a tendency to maintain more abrupt/complete glottal closure. Finally, average LCA increased and LCA variability decreased as the SPL and H_1–H_2 distributions underlying the DPI were classified as more severe, further supporting our interpretation of how LCA muscle activation relates to PVH pathophysiology.

The patients exhibited diminished average CT muscle activation compared to controls, pointing to lower pitches on average. Patients with PVH often voice with reduced pitch variation toward higher pitches, and multiple voice therapy approaches incorporate pitch glides and exaggerated prosody to address this observation (Gartner-Schmidt & Gillespie, 2021; Roy et al., 1997; Verdolini, 2008). Empirically, previous studies identified significantly reduced $F0$ standard deviation in patients' daily life compared to matched controls (Mehta et al., 2015; Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020; Van Stan, Mehta, Ortiz, Burns, Toles, et al., 2020; Van Stan, Ortiz, Marks, et al., 2021). This difference normalized after voice therapy, not after surgical removal of the lesions, insinuating that it may be primarily behavioral. The relationship between mean CT activation and the DPI showed that decreased CT activation related to increased potential for phonotrauma (Van Stan, Mehta, Ortiz, Burns, Marks, et al., 2020; Van Stan, Ortiz, Marks, et al., 2021). This relationship is counterintuitive because traditional vocal dose measures often rely on the opposite relationship: higher cumulative collision forces as $F0$ (i.e., CT activation levels) increases (Rantala & Vilkman, 1999; Titze & Hunter, 2015). However, vocal frequency is known to be motorically redundant as multiple muscles combine to produce different pitches (Titze & Story, 2002). Perhaps voicing at a specific frequency with less CT muscle activation requires laryngeal configurations with higher potential for phonotrauma, that is, increased LCA muscle activation.

There was no notable difference in TA muscle activation between patients and controls, as well as minimal changes in TA activation at high DPI levels. While multiple theories underlying PVH hypothesize general increases in intrinsic laryngeal muscle activation levels, the results

suggest that PVH is an imbalance among these muscles. Specifically, the imbalance may manifest as a "typical" level of TA activation, "higher-than-typical" level of LCA activation, and "lower-than-typical" level of CT activation. Similar imbalances have been found in patients with PVH using superficial electromyography intermuscular coherence between the left and right anterior neck muscles (Stepp et al., 2010, 2011). Future research should explore the imbalances between the left and right activation levels of the same muscle. The results show that the proposed methodology, estimate using just SPL and H_1–H_2 distributions, fails to constrain the behavior of the TA muscle, beyond the uniform distribution of the synthetic data and the prevalence of more simulations with high activation. Therefore, further studies will focus on developing new features that more accurately correlate with its activity. This will lead to more reliable estimates and a deeper understanding of its role in hyperfunction.

Patients exhibited increased average P_S compared to the controls, likely reflecting the increased effort/forces necessary to voice with excessive phonotrauma and/or in the presence of phonotraumatic lesions. Previous work has shown that, compared to controls, patients spend more time at higher amplitudes—that is, increased negative skewness of the neck skin acceleration amplitude. Perhaps higher overall average P_Ss are physiologically necessary to maintain a negatively skewed vocal intensity behavior in daily life. Supporting this interpretation, there was a direct relation between increases in DPI values and increases in P_S.

There are multiple limitations of this study. We acknowledge that the model simulations are bound by a certain level of granularity, as evidenced by a 77% prediction accuracy for SPL and H_1–H_2 values. One potential way to increase the prediction accuracy in future studies includes skipping the transformation of the ambulatory ACC signal into the model glottal airflow signal using the IBIF. While the IBIF transformation allows the ACC to be included in the model, it can introduce noise into the feature calculation (Cortés, Alzamendi, et al., 2022; Morales et al., 2023). However, the ACC has been shown to exhibit a strong relationship with airflow features in previous research (Espinoza et al., 2020; Mehta et al., 2019), justifying its inclusion, even with the use of IBIF. Future improvements could explore alternative methods for incorporating the ACC signal into the model, potentially by developing a dedicated model for ACC within the overall phonation model framework. Refining the model resolution could potentially improve its predictive fidelity. Despite this shortcoming, our model shows promising generalizability, simulating vocal features across a varied population with a fair degree of success. The complexity of inverse mapping—deriving muscle activations and pressures from SPL and H_1–H_2 readings—remains a challenging endeavor due to the nonuniqueness of the solution. Our model currently mitigates this challenge through a stochastic process in sample selection. In this study, we aim to delineate the scope of the DPI space estimation and demonstrate how it characterizes PVH. Looking forward, we aim to enhance our methodology by integrating more measured variables to tighten the range of potential solutions and by expanding our simulation library. Future research efforts will be directed toward adopting more sophisticated approaches, including interpolation methods, and leveraging machine learning techniques such as neural networks, to refine parameter estimation and tackle the issue of nonuniqueness with greater precision.

Conclusions

This study contributes to voice assessment by incorporating numerical modeling techniques and an inverse mapping technique that noninvasively estimate difficult-to-measure physiological parameters such as P_S and intrinsic muscle activation levels. This approach has been further validated by correlating it to measures underlying the DPI, as well as changes in DPI severity. Shifts in P_S, LCA muscle activity, and CT muscle activity during voiced speech in daily life can provide novel insights into the biomechanical responses associated with PVH.

Data Availability Statement

De-identified overall average values for variables in this study are included as supplemental materials. Because the data are from human subjects, more detailed data may be available upon request, if appropriate, and require a data use agreement. Anyone wishing to request access to the data must first contact Sarah DeRosa, Program Coordinator for Research and Clinical Speech-Language Pathology, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, at sederosa@partners.org.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant P50DC015446 (PI: Robert E. Hillman); Agencia Nacional de Investigación y Desarrollo/Becas de Doctorado Nacional Grants 21202490 (PI: Jesús A. Parra) and 21190074 (PI: Emiro J. Ibarra), FONDECYT 1230828 and BASAL AFB240002 (PI for both: Matías Zañartu); and Universidad Técnica Federico Santa María Grant DPP PIIC No. 042/

2021 (PI: Jesús A. Parra). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Alzamendi, G. A., Manríquez, R., Hadwin, P. J., Deng, J. J., Peterson, S. D., Erath, B. D., Mehta, D. D., Hillman, R. E., & Zañartu, M. (2020). Bayesian estimation of vocal function measures using laryngeal high-speed videoendoscopy and glottal airflow estimates: An *in vivo* case study. *The Journal of the Acoustical Society of America, 147*(5), Article EL434. https://doi.org/10.1121/10.0001276

Alzamendi, G. A., Peterson, S. D., Erath, B. D., Hillman, R. E., & Zañartu, M. (2022). Triangular body-cover model of the vocal folds with coordinated activation of the five intrinsic laryngeal muscles. *The Journal of the Acoustical Society of America, 151*(1), 17–30. https://doi.org/10.1121/10.0009169

Behrman, A., Rutledge, J., Hembree, A., & Sheridan, S. (2008). Vocal hygiene education, voice production therapy, and the role of patient adherence: A treatment effectiveness study in women with phonotrauma. *Journal of Speech, Language, and Hearing Research, 51*(2), 350–366. https://doi.org/10.1044/1092-4388(2008/026)

Bhattacharya, P., & Siegmund, T. (2015). The role of glottal surface adhesion on vocal folds biomechanics. *Biomechanics and Modeling in Mechanobiology, 14*(2), 283–295. https://doi.org/10.1007/s10237-014-0603-7

Calvache, C., Solaque, L., Velasco, A., & Peñuela, L. (2023). Biomechanical models to represent vocal physiology: A systematic review. *Journal of Voice, 37*(3), 465.e1–465.e18. https://doi.org/10.1016/j.jvoice.2021.02.014

Carullo, A., Vallan, A., & Astolfi, A. (2013). Design issues for a portable vocal analyzer. *IEEE Transactions on Instrumentation and Measurement, 62*(5), 1084–1093. https://doi.org/10.1109/TIM.2012.2236724

Cheyne, H. A., Hanson, H. M., Genereux, R. P., Stevens, K. N., & Hillman, R. E. (2003). Development and testing of a portable vocal accumulator. *Journal of Speech, Language, and Hearing Research, 46*(6), 1457–1467. https://doi.org/10.1044/1092-4388(2003/113)

Chhetri, D. K., Neubauer, J., Sofer, E., & Berry, D. A. (2014). Influence and interactions of laryngeal adductors and cricothyroid muscles on fundamental frequency and glottal posture control. *The Journal of the Acoustical Society of America, 135*(4), 2052–2064. https://doi.org/10.1121/1.4865918

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum.

Cortés, J. P., Alzamendi, G. A., Weinstein, A. J., Yuz, J. I., Espinoza, V. M., Mehta, D. D., Hillman, R. E., & Zañartu, M. (2022). Kalman filter implementation of subglottal impedance-based inverse filtering to estimate glottal airflow during phonation. *Applied Sciences, 12*(1), Article 401. https://doi.org/10.3390/app12010401

Cortés, J. P., Espinoza, V. M., Ghassemi, M., Mehta, D. D., Van Stan, J. H., Hillman, R. E., Guttag, J. V., & Matías, Z. (2018). Ambulatory assessment of phonotraumatic vocal hyperfunction using glottal airflow measures estimated from neck-surface acceleration. *PLOS ONE, 13*(12), Article e0209017. https://doi.org/10.1371/journal.pone.0209017

Cortés, J. P., Lin, J. Z., Marks, K. L., Espinoza, V. M., Ibarra, E. J., Zañartu, M., Hillman, R. E., & Mehta, D. D. (2022). Ambulatory monitoring of subglottal pressure estimated from neck-surface vibration in individuals with and without voice disorders. *Applied Sciences, 12*(21), Article 10692. https://doi.org/10.3390/app122110692

Dikkers, F. G., & Nikkels, P. G. (1995). Benign lesions of the vocal folds: Histopathology and phonotrauma. *Annals of Otology, Rhinology & Laryngology, 104*(9), 698–703. https://doi.org/10.1177/000348949510400905

Döllinger, M., Zhang, Z., Schoder, S., Šidlof, P., Tur, B., & Kniesburges, S. (2023). Overview on state-of-the-art numerical modeling of the phonation process. *Acta Acustica, 7,* Article 25. https://doi.org/10.1051/aacus/2023014

Erath, B. D., Zañartu, M., & Peterson, S. D. (2017). Modeling viscous dissipation during vocal fold contact: The influence of tissue viscosity and thickness with implications for hydration. *Biomechanics and Modeling in Mechanobiology, 16*(3), 947–960. https://doi.org/10.1007/s10237-016-0863-5

Erath, B. D., Zañartu, M., Stewart, K. C., Plesniak, M. W., Sommer, D. E., & Peterson, S. D. (2013). A review of lumped-element models of voiced speech. *Speech Communication, 55*(5), 667–690. https://doi.org/10.1016/j.specom.2013.02.002

Espinoza, V. M., Mehta, D. D., Van Stan, J. H., Hillman, R. E., & Zañartu, M. (2020). Glottal aerodynamics estimated from neck-surface vibration in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 63*(9), 2861–2869. https://doi.org/10.1044/2020_JSLHR-20-00189

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

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

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

Gartner-Schmidt, J., & Gillespie, A. I. (2021). Conversation training therapy: Let's talk it through. *Seminars in Speech and Language, 42*(1), 32–40. https://doi.org/10.1055/s-0040-1722751

Gomez, P., Schutzenberger, A., Semmler, M., & Dollinger, M. (2018). Laryngeal pressure estimation with a recurrent neural network. *IEEE Journal of Translational Engineering in Health and Medicine, 7,* Article 2000111. https://doi.org/10.1109/JTEHM.2018.2886021

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, Language, and Hearing Research, 32*(2), 373–392. https://doi.org/10.1044/jshr.3202.373

Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zañartu, M., & Mehta, D. D. (2020). An updated theoretical framework for vocal hyperfunction. *American Journal of Speech-Language Pathology, 29*(4), 2254–2260. https://doi.org/10.1044/2020_AJSLP-20-00104

Ibarra, E. J., Parra, J. A., Alzamendi, G. A., Cortés, J. P., Espinoza, V. M., Mehta, D. D., Hillman, R. E., & Zañartu, M. (2021). Estimation of subglottal pressure, vocal fold collision pressure, and intrinsic laryngeal muscle activation from

neck-surface vibration using a neural network framework and a voice production model. *Frontiers in Physiology, 12,* Article 732244. https://doi.org/10.3389/fphys.2021.732244

Lammert, A. C., Melot, J., Sturim, D. E., Hannon, D. J., DeLaura, R., Williamson, J. R., Ciccarelli, G., & Quatieri, T. F. (2020). Analysis of phonetic balance in standard English passages. *Journal of Speech, Language, and Hearing Research, 63*(4), 917–930. https://doi.org/10.1044/2020_JSLHR-19-00001

Lin, J. Z., Espinoza, V. M., Marks, K. L., Zañartu, M., & Mehta, D. D. (2020). 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

Marks, K. L., Lin, J. Z., Burns, J. A., Hron, T. A., Hillman, R. E., & Mehta, D. D. (2020). Estimation of subglottal pressure from neck surface vibration in patients with voice disorders. *Journal of Speech, Language, and Hearing Research, 63*(7), 2202–2218. https://doi.org/10.1044/2020_JSLHR-19-00409

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

Mehta, D. D., Espinoza, V. M., Van Stan, J. H., Zañartu, M., & Hillman, R. E. (2019). The difference between first and second harmonic amplitudes correlates between glottal airflow and neck-surface accelerometer signals during phonation. *The Journal of the Acoustical Society of America, 145*(5), EL386–EL392. https://doi.org/10.1121/1.5100909

Mehta, D. D., Kobler, J. B., Zeitels, S. M., Zañartu, M., Ibarra, E. J., Alzamendi, G. A., Manriquez, R., Erath, B. D., Peterson, S. D., Petrillo, R. H., & Hillman, R. E. (2021). Direct measurement and modeling of intraglottal, subglottal, and vocal fold collision pressures during phonation in an individual with a hemilaryngectomy. *Applied Sciences, 11*(16), Article 7256. https://doi.org/10.3390/app11167256

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,* Article 155. https://doi.org/10.3389/fbioe.2015.00155

Mehta, D. D., Zañartu, M., Feng, S. W., Cheyne, H. A., II., & Hillman, R. E. (2012). Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Transactions on Biomedical Engineering, 59*(11), 3090–3096. https://doi.org/10.1109/TBME.2012.2207896

Morales, A., Yuz, J. I., Cortés, J. P., Fontanet, J. G., & Zañartu, M. (2023). Glottal airflow estimation using neck surface acceleration and low-order Kalman smoothing. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31,* 2055–2066. https://doi.org/10.1109/TASLP.2023.3277269

Nudelman, C. J., Ortiz, A. J., Fox, A. B., Mehta, D. D., Hillman, R. E., & Van Stan, J. H. (2022). Daily Phonotrauma Index: An objective indicator of large differences in self-reported vocal status in the daily life of females with phonotraumatic vocal hyperfunction. *American Journal of Speech-Language Pathology, 31*(3), 1412–1423. https://doi.org/10.1044/2022_AJSLP-21-00285

Parra, J. A., Calvache, C., & Zañartu, M. (2023). Assessing phoneme distribution for speech modeling. *18th International Symposium on Medical Information Processing and Analysis, Valparaiso, Chile, 12567,* 366–372. https://doi.org/10.1117/12.2670042

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*(2), 695–698. https://doi.org/10.1121/1.410307

Popolo, P. S., Švec, J. G., & Titze, I. R. (2005). Adaptation of a pocket PC for use as a wearable voice dosimeter. *Journal of Speech, Language, and Hearing Research, 48*(4), 780–791. https://doi.org/10.1044/1092-4388(2005/054)

Rantala, L., & Vilkman, E. (1999). Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *Journal of Voice, 13*(4), 484–495. https://doi.org/10.1016/S0892-1997(99)80004-6

Roy, N., Bless, D. M., Heisey, D., & Ford, C. N. (1997). Manual circumlaryngeal therapy for functional dysphonia: An evaluation of short- and long-term treatment outcomes. *Journal of Voice, 11*(3), 321–331. https://doi.org/10.1016/S0892-1997(97)80011-2

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

Searl, J., & Dietsch, A. (2014). Testing of the VocaLog vocal monitor. *Journal of Voice, 28*(4), 523.e27–523.e37. https://doi.org/10.1016/j.jvoice.2014.01.009

Sepúlveda, J., Parra, J., Ibarra, E., Araya, M., Cuadra, P. D. L., & Zañartu, M. (2024). Estimation of physiological vocal features from neck surface acceleration signals using probabilistic Bayesian neural networks. *TechRxiv*. https://doi.org/10.36227/techrxiv.172555517.72063796/v1

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2010). Use of neck strap muscle intermuscular coherence as an indicator of vocal hyperfunction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18*(3), 329–335. https://doi.org/10.1109/TNSRE.2009.2039605

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2011). Modulation of neck intermuscular Beta coherence during voice and speech production. *Journal of Speech, Language, and Hearing Research, 54*(3), 836–844. https://doi.org/10.1044/1092-4388(2010/10-0139)

Story, B. H. (2008). Comparison of magnetic resonance imaging–based vocal tract area functions obtained from the same speaker in 1994 and 2002. *The Journal of the Acoustical Society of America, 123*(1), 327–335. https://doi.org/10.1121/1.2805683

Story, B. H., & Titze, I. R. (1995). Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America, 97*(2), 1249–1260. https://doi.org/10.1121/1.412234

Titze, I. R., & Hunter, E. J. (2007). A two-dimensional biomechanical model of vocal fold posturing. *The Journal of the Acoustical Society of America, 121*(4), 2254–2260. https://doi.org/10.1121/1.2697573

Titze, I. R., & Hunter, E. J. (2015). Comparison of vocal vibration-dose measures for potential-damage risk criteria. *Journal of Speech, Language, and Hearing Research, 58*(5), 1425–1439. https://doi.org/10.1044/2015_JSLHR-S-13-0128

Titze, I. R., & Story, B. H. (2002). Rules for controlling low-dimensional vocal fold models with muscle activation. *The Journal of the Acoustical Society of America, 112*(3), 1064–1076. https://doi.org/10.1121/1.1496080

Titze, I. R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America, 91*(5), 2936–2946. https://doi.org/10.1121/1.402929

Vahabzadeh-Hagh, A. M., Zhang, Z., & Chhetri, D. K. (2018). Hirano's cover-body model and its unique laryngeal postures revisited. *The Laryngoscope, 128*(6), 1412–1418. https://doi.org/10.1002/lary.27000

Van Stan, J. H., Mehta, D. D., Ortiz, A. J., Burns, J. A., Marks, K. L., Toles, L. E., Stadelman-Cohen, T., Krusemark, C., Muise, J., Hron, T., Zeitels, S. M., Fox, A. B., & Hillman, R. E. (2020). Changes in a Daily Phonotrauma Index after laryngeal surgery and voice therapy: Implications for the role of daily voice use in the etiology and pathophysiology of phonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 63*(12), 3934–3944. https://doi.org/10.1044/2020_JSLHR-20-00168

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

Van Stan, J. H., Ortiz, A. J., Cortes, J. P., Marks, K. L., Toles, L. E., Mehta, D. D., Burns, J. A., Hron, T., Stadelman-Cohen, T., Krusemark, C., Muise, J., Fox-Galalis, A. B., Nudelman, C., Zeitels, S., & Hillman, R. E. (2021). Differences in daily voice use measures between female patients with nonphonotraumatic vocal hyperfunction and matched controls. *Journal of Speech, Language, and Hearing Research, 64*(5), 1457–1470. https://doi.org/10.1044/2021_JSLHR-20-00538

Van Stan, J. H., Ortiz, A. J., Marks, K. L., Toles, L. E., Mehta, D. D., Burns, J. A., Hron, T., Stadelman-Cohen, T., Krusemark, C., Muise, J., Fox, A. B., Nudelman, C., Zeitels, S., & Hillman, R. E. (2021). Changes in the Daily Phonotrauma Index following the use of voice therapy as the sole treatment for phonotraumatic vocal hyperfunction in females. *Journal of Speech, Language, and Hearing Research, 64*(9), 3446–3455. https://doi.org/10.1044/2021_JSLHR-21-00082

Verdolini, K. (2008). *Lessac–Madsen resonant voice therapy: Clinician manual*. Plural Publishing.

Verdolini, K., Druker, D. G., Palmer, P. M., & Samawi, H. (1998). Laryngeal adduction in resonant voice. *Journal of Voice, 12*(3), 315–327. https://doi.org/10.1016/S0892-1997(98)80021-0

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

Zhang, Z. (2020). Estimation of vocal fold physiology from voice acoustics using machine learning. *The Journal of the Acoustical Society of America, 147*(3), EL264–EL270. https://doi.org/10.1121/10.0000927

Zhang, Z. (2022). Estimating subglottal pressure and vocal fold adduction from the produced voice in a single-subject study (L). *The Journal of the Acoustical Society of America, 151*(2), 1337–1340. https://doi.org/10.1121/10.0009616