ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Characterization of Vocal Motor Control Using Laryngeal Kinematics in Individuals With Hyperfunctional Voice Disorders

Hasini R. Weerathunge,[a,b] Jenny Vojtech,[b] Courtney J. Dunsmuir,[b] Sarah J. Cocroft,[b] Manuel E. Díaz-Cádiz,[b] Victoria McKenna,[b,c] and Cara E. Stepp[a,b,d]

[a] Department of Biomedical Engineering, Boston University, MA [b] Department of Speech, Language, and Hearing Sciences, Boston University, MA [c] Department of Communication Sciences and Disorders, University of Central Florida, Orlando [d] Department of Otolaryngology–Head and Neck Surgery, Boston University School of Medicine, MA

ARTICLE INFO

Article History:
Received August 28, 2024
Revision received November 6, 2024
Accepted January 1, 2025

Editor-in-Chief: Rita R. Patel
Editor: Nancy Pearl Solomon

https://doi.org/10.1044/2025_JSLHR-24-00598

ABSTRACT

Purpose: High-speed videoendoscopy was used to investigate how underlying laryngeal motor control strategies differ in individuals with and without hyperfunctional voice disorders (HVDs). Three laryngeal kinematic measures were defined to characterize laryngeal motor control: kinematic stiffness, spatiotemporal index, and asymmetry index.

Method: Twenty-eight adults with HVDs and 28 age- and sex-matched controls produced repeated utterances of /ifi/ at three different gesture rates (50, 65, and 80 beats per minute) and three self-induced vocal effort levels (mild, moderate, and maximum effort) to elicit a range of linguistic contexts for the vocal targets produced. The glottal angle profiles of /ifi/ productions were extracted to calculate three kinematic measures of laryngeal motor control: kinematic stiffness (estimating laryngeal muscle tension), spatiotemporal index (estimating production variability), and asymmetry index (estimating movement asymmetry).

Results: Individuals with HVDs exhibited statistically significantly higher kinematic stiffness during varying effort levels and higher spatiotemporal indices and asymmetry indices compared to controls, indicating higher laryngeal muscle tension, production variability, and movement asymmetry, respectively.

Conclusion: Laryngeal kinematics suggest differing underlying motor control strategies in individuals with HVD relative to controls, which may inform better understanding of the etiology of HVDs.

Supplemental Material: https://doi.org/10.23641/asha.28550387

Approximately 10% of the adult population in the United States suffers from voice issues (Bhattacharyya, 2014), and 40% of this population present with a hyperfunctional component (Hillman et al., 2020). Hyperfunctional voice disorders (HVDs) are conditions associated with excessive perilaryngeal musculoskeletal activity during phonation, also known as vocal hyperfunction (VH; Oates & Winkworth, 2008). HVDs are characterized by altered vocal quality, pitch, and loudness (Hillman et al., 1989). Individuals with HVD also report increases in self-perceived vocal effort (Altman et al., 2005; Roy et al., 2004). Although VH is widely prevalent in clinical practice, effective treatment of voice disorders associated with VH is impeded by limited knowledge of its underlying etiology and pathophysiology. The VH framework proposed by Hillman et al. (1989) and updated in 2020 suggests that a cycle of vocal misuse causes individuals to adopt compensatory hyperfunctional vocal patterns (characterized by increased laryngeal muscle tension) that may persist even when the causes that elicited VH are no longer present. Two main subtypes of presentation have been identified. VH in the absence of structural dysfunction is termed "nonphonotraumatic VH" (also known as muscle tension dysphonia), whereas benign phonotrauma that arises and persists due to hyperfunctional vocal behavior (e.g., vocal fold nodules and polyps) is termed "phonotraumatic VH."

Correspondence to Cara E. Stepp: cstepp@bu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

However, this framework does not explain how two individuals with the same initial symptom presentation (e.g., an upper respiratory infection) could develop vastly different vocal patterns (i.e., either recovering completely or chronically developing VH with symptoms of dysphonia or laryngeal muscle tension). The onset and chronic persistence of HVDs have been attributed to a variety of etiological factors including daily voice use and vocal patterns (Altman et al., 2005; Van Houtte et al., 2011), anatomical and/or physiological vulnerability (Mi et al., 2010), psychological and/or personality predisposition (Baker & Ben-Tovim, 1996; Dietrich et al., 2008; Hillman et al., 1989; Roy et al., 2000), altered biomechanics (Abur et al., 2022; McKenna et al., 2020; Stepp et al., 2017; Tam et al., 2018; Ziethe et al., 2019), and sensorimotor deficits (Abur et al., 2022; McKenna et al., 2020; Stepp et al., 2017; Tam et al., 2018; Ziethe et al., 2019). However, the underlying pathophysiology of the disease remains less understood, in particular how laryngeal motor control may be altered in individuals with HVDs.

Laryngeal Motor Control in Individuals With HVDs

Recent studies examining HVDs implicate atypical laryngeal motor control as a contributing factor to hyperfunctional behaviors. Research using altered auditory feedback paradigms suggests that individuals with HVDs have difficulty using auditory feedback to update their feedforward control subsystem (Abur et al., 2021; Nguyen et al., 2022; Tam et al., 2018). Additional findings suggest possible auditory–perceptual deficits and, moreover, larger individual variability in vocal responses (Abur et al., 2021; Stepp et al., 2017). However, these studies also showcase the inherent difficulties in carrying out auditory feedback perturbation studies in this population. Individuals with HVDs may be unable to sustain vowels for longer periods of time, and there may be difficulty tracking vocal fundamental frequency in dysphonia (as the level of aperiodicity in a speech signal increases when voice problems are present; Eadie & Doyle, 2005; Titze, 1991). Thus, it is crucial to identify additional methods to characterize laryngeal motor control in this population.

An alternative way to study laryngeal motor control that does not rely on periodic speech acoustics is by observing the motion of laryngeal articulators (i.e., laryngeal kinematics) directly. Laryngeal kinematics have not been studied extensively due to the relative inaccessibility of intrinsic laryngeal structures and the limited sampling rates of traditional videoendoscopy (i.e., 30 frames per second [fps]; Mehta & Hillman, 2012). A standard 30-fps sampling rate can capture an average of three to seven frames during posturing movements of the laryngeal articulators, such as vocal fold abduction and adduction. In contrast, Diaz-Cadiz et al. (2018) noted that a frame rate of at least 120–150 fps is required to sufficiently sample vocal fold trajectories. With the introduction of high-speed videoendoscopy (HSV)—which can be used to capture images of the vocal folds at ≥ 1,000 fps—there is potential to study peripheral laryngeal motor control in individuals with voice disorders via laryngeal kinematics with sufficient spatial and temporal resolution (Freeman et al., 2012; Iwahashi et al., 2016). The focus of this study was to use HSV to explore laryngeal motor control via three aspects of laryngeal kinematics: laryngeal muscle tension, production variability, and movement asymmetry.

Laryngeal Muscle Tension

Due to altered biomechanics and/or underlying neurophysiology, individuals with HVDs are thought to have elevated levels of tension in their intrinsic laryngeal structures (Hillman et al., 1989, 2020). Unfortunately, directly quantifying tension is challenging due to a lack of techniques that can measure tension in vivo in an ecologically valid manner.

Laryngeal kinematics may be a feasible way to estimate laryngeal muscle tension in individuals with HVDs. Kinematic stiffness (KS) is a biomechanical correlate of muscular tension that has been used in limb movement literature (J. Cooke, 1980). It is an estimate of resistance to movement, defined as the ratio between maximum velocity of movement to the maximum movement extent (Shiller et al., 2002). KS was eventually adapted to assess articulatory movements (Munhall et al., 1985; Ostry et al., 1987) and, more recently, laryngeal movements (A. Cooke et al., 1997; Stepp et al., 2010).

Stepp et al. (2010) investigated KS during vocal fold abduction and adduction using a biomechanical model of laryngeal dynamics. The model predicted that increased KS could reflect increased laryngeal muscle tension. The authors found that increases in gesture rate (thought to increase global laryngeal muscle tension) did increase KS values. However, these increases were greater in individuals without HVDs relative to those with HVDs, a finding interpreted by the authors as signifying increased *baseline* laryngeal muscle tension present in individuals with HVDs. One technical limitation of Stepp et al. (2010) was the sparse sampling rate of conventional videoendoscopy (30 fps), which required a function to be fitted to estimate KS. HSV can be used to overcome possible estimation errors associated with function fits of KS to more accurately characterize laryngeal KS—and thereby, tension.

Production Variability

Speech production variability is increased in HVDs, likely due to deficiencies in updating predefined motor commands (Abur et al., 2021; Stepp et al., 2017). This variability is frequently noted in clinical reports of HVDs and is backed by research findings that demonstrate variability in aerodynamics (e.g., phonatory airflow; Belsky et al., 2021; Gillespie et al., 2013; Higgins et al., 1999), acoustic measures of voice onset time (McKenna et al., 2020) and mean voice intensity (Belsky et al., 2021), tongue kinematics (Shipurkar et al., 2023), and sensorimotor measures (Abur et al., 2021). Using a measure of laryngeal kinematics that can characterize both spatial and temporal production variability may provide more insight into production variability in this population.

The spatiotemporal index (STI; Smith et al., 2000) has been proposed as such a measure and has been used extensively to characterize articulatory kinematics. STI quantifies the consistency of movement patterns—such as speech gestures—by examining both spatial and temporal variability across multiple repetitions of the same target. Each signal is amplitude-normalized using *z* scores to focus on pattern consistency rather than individual signal features (e.g., vocal loudness), and all signals are time-aligned to a fixed length to control for general duration differences. Standard deviations are calculated across repetitions as a function of the normalized time duration. Summing these deviations yields the STI score between 0 and approximately 50 (based on the number of repetitions; see Wisler et al., 2022, for a comprehensive overview of the fundamental limits of STI), where lower values indicate greater movement stability and reduced variability across repetitions. The STI has not previously been applied to laryngeal kinematics, but we anticipate that it may be sensitive to differences in vocal motor strategies that may differentiate individuals with and without HVDs.

Movement Asymmetry

Motor invariance theories suggest that there is a predefined motor command for each speech target. Altering speaking conditions, such as gesture rate or vocal effort level, therefore leads to adjustments in covariable parameters such as movement duration and maximum movement velocity (Gracco & Abbs, 1986); this has been observed empirically in articulatory kinematics of the lip, tongue tip, and jaw (Munhall, 1984). These adjustments suggest a common underlying sensorimotor mechanism with hard-coded temporal scaling.

Leveraging this theory of motor invariance to examine laryngeal kinematics may offer insights into how sensorimotor control is coded at the laryngeal level. Tracking movement trajectories using *velocity profiles* (Munhall et al., 1985) is a common approach in articulatory motor control. Velocity profile asymmetry, defined as the ratio between acceleration duration and deceleration duration, serves as an index of motor control asymmetry. Studies have shown that this asymmetry varies based on the type of articulatory movement (e.g., tongue dorsum vs. tip, jaw, low lip), with gestures to open and close the mouth demonstrating differential effects (e.g., jaw lowering or elevating; Adams, 1990; Adams et al., 1993; Connor et al., 1988; Ostry et al., 1987). It is postulated that different asymmetry patterns exist based on articulator type due to different synergies required for efficient speech production (Gracco, 1988). Relatedly, velocity profile asymmetries vary as a function of overall movement speed, with greater asymmetry during slower movements than faster movements (Beggs & Howarth, 1972). Modeling work by Bullock and Grossberg (1988) suggest that this increase in velocity profile asymmetries during longer movement durations may indicate less reliance on open-loop control (in which preplanned movement commands constitute movement via feedforward mechanisms) relative to closed-loop control (in which online modifications to movement commands are made via feedback-based corrections).

Building on this knowledge, examining the asymmetry of laryngeal velocity profiles across different speaking conditions—such as different gesture rates or levels of self-induced vocal effort—may provide insight into the differential contributions of feedback versus feedforward laryngeal motor control strategies. To understand the asymmetries of laryngeal motor control, we can make use of an asymmetry index (AI) originally proposed by Ostry et al. (1987) to characterize the velocity profile of articulatory gestures; here, we define AI as the ratio between the deceleration and acceleration durations. As vocal fold adduction is a closing gesture in which the vocal folds approximate, we would expect time spent in deceleration phase to increase as gesture rate slows (and vice versa).

Although velocity profiles have been extensively studied for articulatory kinematics, laryngeal velocity profiles (particularly for vocal fold approximation) remain relatively unexplored. Thus, considering potential deficits in sensorimotor integration in individuals with HVDs, examining the asymmetry of laryngeal velocity profiles may offer valuable insights into feedforward control under different speaking conditions.

Research Statement

In this study, we investigated whether underlying laryngeal motor control strategies differ between individuals with and without HVDs. We quantitatively characterized laryngeal motor control via KS for laryngeal muscle tension, STI for vocal production variability, and AI for movement asymmetry. We assessed these kinematic measures across different speaking conditions—including three gesture rates (slow, medium, fast) and three self-induced levels of vocal effort (mild, moderate, maximum)—for ecological validity of produced vocal targets. Our primary hypothesis was that individuals with HVDs would exhibit significantly increased KS, STI, and AI values compared to controls. We anticipated that increased gesture rates or self-induced levels of vocal effort would result in changes in laryngeal motor strategies that would have effects on KS, STI, and AI, but that these changes might be mitigated in individuals with HVDs due to their higher baseline values of KS, STI, and AI (i.e., a ceiling effect).

Method

Participants

HSV data collected from a group of 28 adults with HVDs (cisgender female = 23, cisgender male = 4, nonbinary = 1) aged 19–70 years (*M* = 37.5 years, *SD* = 16.1 years) and 28 age- and sex-matched controls (female = 23, male = 5) aged 18–68 years (*M* = 34.4 years, *SD* = 17.8 years) were selected from an existing database collected from 2017 to 2020 (McKenna et al., 2016; Vojtech & Stepp, 2024).

Diagnosis in individuals with HVDs was based on the medical judgment of one of four referring board-certified laryngologists with access to their own visual and auditory examination and patient-reported symptoms. Criteria for diagnosis in individuals considered to have a nonphonotraumatic voice disorder included increased supraglottic compression (judged via laryngeal stroboscopy), increased paralaryngeal muscle tension during palpation, and the absence of neurological and organic findings. The control group comprised adults without history of speech, language, hearing, neurological, or voice disorders. All data were collected with informed consent obtained prior to participation from all participants in compliance with the Boston University Institutional Review Board (Protocol #2625).

All participants were screened by a certified voice-specializing speech-language pathologist for vocal quality via the 100-mm visual analog scale for overall severity of dysphonia from the Consensus Auditory–Perceptual Evaluation of Voice (CAPE-V) with nonlinearly placed textual labels for severity as originally published (American Speech-Language-Hearing Association, 2002) and flexible nasoendoscopy laryngeal imaging. Overall severity of dysphonia ratings ranged from 0.9 to 39.0 mm (*M* = 12.1, *SD* = 7.0) in the HVD group and from 0.6 to 23.5 mm (*M* = 7.0, *SD* = 5.1) in the control group. The speech-language pathologist re-rated a random selection of 13 of 56 participants (23%) in a separate session to assess intrarater reliability. We calculated two-way, mixed-effects intraclass correlation coefficients (ICCs) based on the absolute agreement of single measures and observed good reliability, with an ICC of .86. Detailed demographics of the two groups are reported in Table 1.

Recording Procedure

Task Training

Participants received training to produce eight consecutive iterations of the vowel–consonant–vowel (VCV) utterance, /ifi/. The use of /ifi/ as a production provided clear vocal fold adductory gestures at the end of the consonant (/f/). Subsequently, individuals were trained to produce /ifi/ strings at varying speeds (in beats per minute [BPM]) and self-induced levels of vocal effort to elicit an ecologically valid range of conditions for the same VCV target productions (McKenna et al., 2016). A metronome was used to train vocal speeds at a slow rate (50 BPM), a medium rate (65 BPM), and a fast rate (80 BPM). Participants were instructed to increase their vocal effort during speech as if they were trying to push air out while maintaining comfortable pitch and volume. Emphasis on maintaining a comfortable speaking volume was included since vocal sound pressure level (SPL) shows diverging relationships with subglottal pressure and listener-perceived estimates of vocal effort when comparing typical speakers (Rosenthal et al., 2014) and speakers with HVDs (Espinoza et al., 2017; Stepp et al., 2012). Mild effort was described as “mildly more effort than regular speaking voice,” moderate effort as “more effort than mild,” and maximum effort as “as much effort as you can while still having a voice.” Participants altered their vocal effort while maintaining their typical speaking rate. This resulted

Table 1. Demographics of individuals with and without hyperfunctional voice disorders.

	Sex		Age (years)			Diagnosis	
Group	**Female**	**Male**	***M***	***SD***	**Range**	**Subtype**	***n***
HVD	23	5	37.5	16.1	19–70	NPVH	20
						PVH	8
Control	23	5	34.4	17.8	18–68	N/A	

Note. PVH cases: nodules (4), scars (1), scars and polyp (1), polyp (1), lesion (1). HVD = hyperfunctional voice disorder; NPVH = nonphonotraumatic vocal hyperfunction; PVH = phonotraumatic vocal hyperfunction; N/A = not applicable.

in a total of six speaking conditions: three gesture rates (slow, medium, fast) and three vocal effort levels (mild, moderate, maximum).

Experimental Protocol

A headset microphone and neck surface accelerometer were placed on the participants. The neck surface accelerometer was placed on the anterior neck superior to the thyroid notch and inferior to the cricoid cartilage using double-sided adhesive. The headset microphone was calibrated with a sound pressure meter (HBK 2250) placed 7 cm from the lips angled toward the mouth, with acoustic excitation provided by an electrolarynx located at the corner of the mouth. All microphone and accelerometer signals were sampled at 44.1 kHz with 16-bit resolution.

A flexible endoscope (Pentax, Model FNL-10RP3, 3.5-mm) was used to record HSV images at a frame rate of 1,000 fps. For cases in which participant anatomy or comfort interfered with image acquisition, a thinner flexible endoscope (Pentax, Model FNL-7RP3, 2.4 mm) was used (38 of 56 participants). The endoscope was passed transnasally over the soft palate into the hypopharynx for laryngeal visualization. Individuals were instructed to produce the eight-/ifi/ train for each recording. Participants completed at least two recordings for each of the six speaking conditions, with an average of 15 /ifi/ utterances per condition. All repeated productions were included in the initial analysis. Video images were acquired using Photron Fastcam Viewer software (Version 3.6.6). Recordings were trigged using a custom MATLAB algorithm that automatically time-aligned the video images with the microphone and accelerometer signals.

Data Extraction

A MATLAB-based, semiautomated glottic angle extraction software, described in detail in Diaz-Cadiz et al. (2018), was used to calculate glottic angle waveform from each /ifi/ production (Version 9.3, The MathWorks).

Experimental Fidelity

Quantitative analyses performed to confirm the fidelity of the experimental procedures is discussed in detail in Supplemental Material S1. Participants were instructed to (a) produce three distinct gesture rates (in BPM; slow, medium, fast) across the gesture rate conditions and (b) maintain a constant voice SPL across conditions involving variations in self-induced vocal effort at three effort levels. The analysis revealed that actual gesture rates were significantly higher in the fast condition and lower in the slow condition compared to the medium rate ($p < .001$ for both), supporting fidelity for the first objective (see Supplemental Material S1, Figure S1a, and Table S1). Voice SPL did not significantly differ among the effort conditions (mild, moderate, maximum; $p > .05$ for all comparisons), supporting fidelity for the second objective (see Supplemental Material S1, Figure S1b, and Table S2).

Technician Training

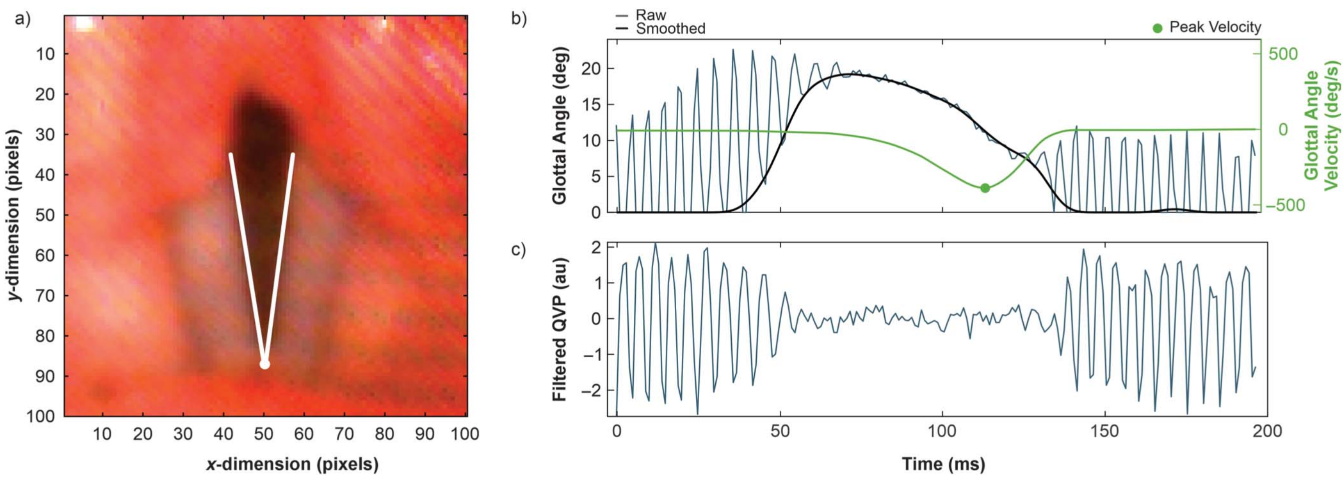
A total of nine research assistants were trained to extract glottal angles using the MATLAB interface from Diaz-Cadiz et al. (2018). Prior to processing experimental data, all technicians were trained in glottal angle (i.e., the angle at anterior commissure made by the two vocal folds) tracking at conventional frame rates of 30 fps. These markings were then compared to a gold-standard marking provided by a trained technician. All nine technicians involved in the angle tracking of the data set were required to have a minimal training standard of .80 via a two-way mixed-effects ICC for consistency of agreement (Diaz-Cadiz et al., 2018). The resulting average reliability for the nine technicians was ICC(3, 1) = .86 (*SD* = .05, range: .82–.95). A second set of four technicians reanalyzed all data to ensure and enhance glottal angle tracking such that movement trajectories of glottal angle and vocal fold adduction velocity were correctly marked to measure the proposed laryngeal kinematic measures of the current study. The resulting average reliability for the four technicians was ICC(3, 1) = .83 (*SD* = .04, range: .80–.88).

Glottal Angle Tracking

Once the technicians acquired the reliability standards, they used the semiautomated glottal angle extraction software to track glottal angles over time of each /ifi/ at 1,000 fps (see Figure 1). Technicians first visually confirmed the usability of /ifi/ image frames, including confirmation that no obstructions were present (e.g., epiglottis or arytenoid cartilages blocking view of the vocal folds), and determined that the automatically estimated glottic angle waveform was appropriate. Usable videos were then processed through the semiautomated glottal angle extraction algorithm, whereas unusable videos were removed from further analysis.

The glottal angle profile (see Figure 1b, gray) that was calculated from the algorithm was smoothed using a 15th-order, zero-phase, low-pass, Hamming window–based finite impulse response filter with a filter cutoff at 25 Hz (Diaz-Cadiz et al., 2018) in preparation for velocity estimation. The 25-Hz cutoff frequency was selected to retain the primary kinematic components of laryngeal articulators during adduction, which typically lasts 104–227 ms (Dailey et al., 2005). The velocity profiles of each /ifi/ utterance were derived by taking the first derivative of these band-limited angle profiles (see Figure 1b, green). Technicians used this information to examine the quality of automated angle tracking. If the angles were not appropriately tracked, the technician was given the option to provide manual markings to guide the algorithm; in the

Figure 1. Semiautomated software for glottal angle extraction for a single /ifi/ production. (a) Videoendoscopic images of vocal folds, with white lines framing the vocal fold edges from the anterior commissure to the vocal processes; (b) raw glottic angle waveform (gray, left axis) with smoothed data overlay (black, left axis) and glottic angle velocity profile (green, right axis) with peak velocity indicated by a filled green circle; and (c) filtered quick vibratory profile (QVP).



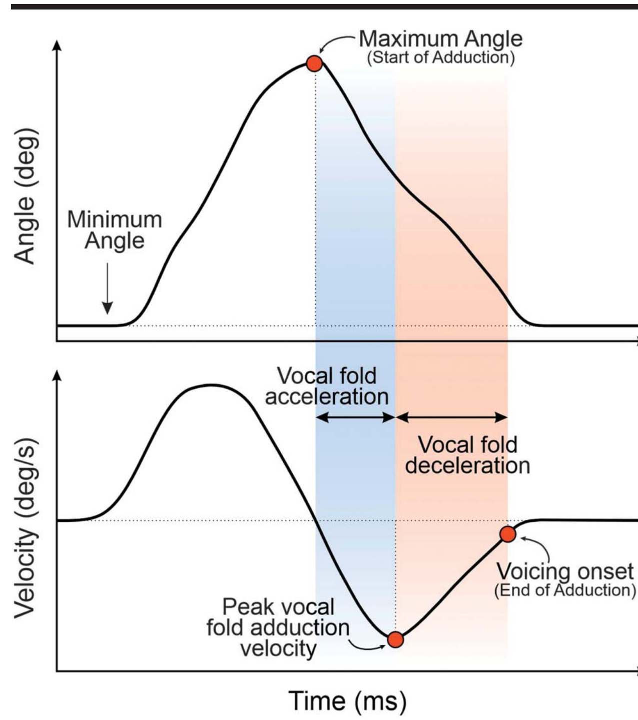
event of persistent tracking errors following manual marking, the technician was given the option to mark the production as unusable.

Vocal Fold Adduction

Once the glottal angle waveform was generated, a secondary analysis was carried out to identify events of vocal fold adduction. Technicians were provided with a slightly modified software showing the time-aligned videos, microphone and accelerometer signals, glottal angle waveform, and a quick vibratory profile (see Figure 1c). As the quick vibratory profile quantifies vocal fold vibratory motion via pixel intensity (Ikuma et al., 2013), it was provided as an alternative to the glottal angle waveform to discriminate the vibrating vocal folds in images of poorer resolution.

Technicians used the provided information to identify a series of data points necessary to characterize the adduction gesture for each /ifi/ production. The start of adduction was defined as the time point of the maximum glottal angle (following the first /i/), and the end was the time point of the first full or maximum vocal fold approximation during voicing onset (see Figure 2, top). Adduction duration was measured as the time between these two points. From the glottal angle velocity waveform, technicians recorded peak adduction velocity as the highest negative velocity (see Figure 2, bottom). This peak velocity divided the adduction gesture into two phases: the acceleration phase (from maximum glottal angle to peak velocity) and the deceleration phase (from peak velocity to voicing onset), during which the vocal folds are approximating. These data points (maximum glottal angle, voicing onset, peak vocal fold adduction velocity, adduction duration, acceleration duration, deceleration duration) were manually marked by experimenters. The

Figure 2. Schematic of vocal fold adduction measures, where maximum and minimum glottal angles are identified from the glottal angle waveform (the minimum determined by a sigmoidal fit as in Diaz-Cadiz et al., 2018). From the glottal angular velocity waveform, peak adduction velocity is the maximum negative velocity, and voicing onset is the point of the first full or maximum vocal fold approximation. Acceleration is measured from the maximum angle to peak velocity, and deceleration is measured from peak velocity to voicing onset, with adduction duration as the sum of these phases.



sigmoidal fit applied to the model was used to derive the minimum glottal angle (see Diaz-Cadiz et al., 2018).

Data Analysis

The extracted glottal angle waveform and single-point measures of vocal fold adduction were then utilized to calculate the three primary kinematic measures: KS, STI, and AI. Note that the glottal angle waveform refers to the time-series trajectory (i.e., time vs. glottal angle) extracted by technicians for an /ifi/ production (see Figure 2a).

- **KS:** The ratio between the peak vocal fold adduction velocity (ω_{peak}) and the maximum extent of the glottic angle during movement, defined as the difference between the maximum (θ_{max}) and minimum (θ_{min}) glottal angles (Equation 1):

$$\text{KS}\ (1/\text{s}) = \frac{\omega_{\text{peak}}\ (\text{deg/s})}{\theta_{\max}\ (\text{deg}) - \theta_{\min}(\text{deg})} \tag{1}$$

- **STI:** The cumulative sum of the standard deviations of vocal fold angles extracted during adduction gesture. To compute STI, we preprocessed each /ifi/ production as follows:

1. Extracted the glottal angle waveform during the adduction gesture (i.e., from the timepoint of maximum glottal angle until voicing onset).
2. Amplitude-normalized the extracted waveform using *z* scores.
3. Aligned the waveform to a fixed 1,000-point vector using cubic interpolation.
4. Sampled the waveform at 2% normalized time intervals, resulting in 50 final data points.

After transforming the adduction gesture of each /ifi/ production into a 50-point normalized angle vector ($\boldsymbol{A}^{\text{norm}}_{\text{adduction}}$), we calculated the standard deviation across all productions at each of the 50 data points. These 50 standard deviations were then summed to yield a single STI value. To address bias due to the number of productions considered, we applied a bias correction factor, γ, based on Wisler et al. (2022), in the final calculation (Equation 2):

$$\text{STI} = \gamma \times \sum SD\left(\boldsymbol{A}^{\text{norm}}_{\text{adduction}}\right) \tag{2}$$

- **AI:** The ratio between the deceleration duration and acceleration duration of the velocity profile (inverse to the measure used in Ostry et al., 1987) is defined as the AI (Equation 3):

$$\text{AI} = \frac{\text{Deceleration Duration (s)}}{\text{Acceleration Duration (s)}} \tag{3}$$

Statistical analyses were completed using Minitab Statistical Software (Version 19; Minitab, Inc.). Two multivariate analysis of variance (MANOVA) models were constructed to examine the laryngeal kinematic measures (KS, STI, AI) according to group and either gesture rate (slow, medium, fast) or vocal effort level (mild, moderate, maximum), including relevant interactions (Group × Gesture Rate or Group × Vocal Effort Level). An alpha level of .05 was used to assess significance for each model. For any significant effects identified in the MANOVAs, post hoc univariate analyses of variance (ANOVAs) were conducted for each dependent variable (KS, STI, AI) to further investigate the source of the effect. Effect sizes (partial eta-squared, η_p^2) were calculated to quantify the magnitude of the effect for each kinematic measure. In an effort to understand if some or all of these kinematic measures assess similar information about the underlying laryngeal motor control mechanisms, we additionally investigated the relationships among them via Spearman's rank correlation coefficient analyses.

Results

Out of the 5,097 total /ifi/ productions, 13.5% of productions were initially discarded prior to processing due to issues related to video usability (n = 688). Technicians accepted the automated results of 57.2% of productions (n = 2,916). Manual angle extraction was performed for 25.0% of productions (n = 1,276). A remaining 4.3% productions were removed after secondary algorithm run due to errors in algorithmic angle tracking that could not be corrected (n = 217). In total, 82.2% of the total /ifi/ productions were used for subsequent analysis (n = 4,192).

Prior to conducting the MANOVAs, the assumptions of multivariate normality and homogeneity of covariance matrices were tested. Box's *M* test confirmed homogeneity of covariance matrices (p = .417 for gesture rate, p = .384 for vocal effort level), and the data were transformed to meet multivariate normality using Box–Cox transformations (p = .990 for gesture rate, p = .153 for vocal effort level; via Shapiro–Wilk test). The dependent variables (KS, STI, AI) were then rescaled using *z*-score normalization before being analyzed in the MANOVA models.

Gesture Rate

The MANOVA on gesture rate revealed a significant multivariate effect of group (Wilks' λ = .942, F(3, 153) = 3.128, p = .028, η_p^2 = .058; see Table 2). Follow-

Table 2. Multivariate analysis of variance results for laryngeal kinematic trajectory measures based on group, gesture rate, and their interaction.

Effect	Wilks' λ	*F*	*df*	*p*	η_p^2	Effect size interpretation
Intercept	.999	0.021	3, 153	.996	.000	Negligible
Group	.942	3.128	3, 153	.028	.058	Small
Rate	.934	1.775	6, 306	.104	.034	Small
Group × Rate	.993	0.186	6, 306	.981	.004	Negligible

Note. Effect size interpretations from Cohen (1988).

up univariate ANOVAs on the Box-Cox–transformed and standardized data, with group as a fixed factor, indicated significant group effects on STI, $F(1, 155) = 7.527$, $p = .007$, $\eta_p^2 = .046$, and AI, $F(1, 155) = 7.969$, $p = .005$, $\eta_p^2 = .049$ (see Table 3), but not on KS. Neither the effect of gesture rate nor its interaction with group significantly influenced the laryngeal kinematic measures ($p > .05$).

As shown in Table 3, analysis of the transformed data showed that the HVD group had significantly higher standardized means compared to the control group for STI (HVD = 0.22, control = −0.21) and AI (HVD = 0.21, control = −0.22). Though not statistically significant, HVD group means were also higher than the control group for KS (HVD = 0.07, control = −0.08). To facilitate interpretation, we back-transformed these values to their original scale, where we observed identical trends. Specifically, the HVD group exhibited larger mean values for KS (HVD = 28.69 1/s, control = 27.19 1/s), STI (HVD = 10.66, control = 8.58), and AI (HVD = 2.29, control = 1.82; see Table 4). Although statistical inferences are based on the transformed data analysis, Figure 3 illustrates the back-transformed estimated marginal means and their 95% confidence intervals to provide a more intuitive understanding of the group differences in their original measurement scales.

Vocal Effort Level

As in the MANOVA on gesture rate, the MANOVA on vocal effort level revealed a significant multivariate effect of group (Wilks' $\lambda = .822$, $F(3, 155) = 11.94$, $p < .001$, $\eta_p^2 = .178$; see Table 5). Follow-up univariate ANOVAs on the Box-Cox–transformed and standardized data, with group as a fixed factor, indicated significant group effects on all dependent variables: KS, $F(1, 157) = 19.896$, $p < .001$, $\eta_p^2 = .112$; STI, $F(1, 157) = 30.451$, $p < .001$, $\eta_p^2 = .162$; and AI, $F(1, 157) = 14.661$, $p < .001$, $\eta_p^2 = .085$ (see Table 6). Neither the effect of vocal effort level nor its interaction with group significantly influenced the laryngeal kinematic measures ($p > .05$). As in the analysis for gesture rate, the HVD group had significantly higher standardized values compared to the control group for KS (HVD = 0.32, control = −0.34), STI (HVD = 0.41, control = −0.39), and AI (HVD = 0.30, control = −0.27; see Table 6). Corresponding trends were observed on the original scale of the dependent variables, wherein individuals with HVD exhibited higher mean values for all dependent variables compared to the control group. Specifically, HVDs had a mean KS of 32.05 1/s (control = 24.98 1/s), STI of 11.81 (control = 7.87), and AI of 2.44 (control = 1.80; see Table 7). Figure 4 presents these back-transformed results for a more intuitive visualization of the group differences on their original measurement scales.

Relationships Among Laryngeal Kinematic Measures

Figure 5 shows the results of the post hoc Spearman's rank correlations for the three laryngeal kinematic measures according to (a–c) gesture rate and (d–f) vocal effort level. For gesture rate, the relationship between KS and STI was significant and strong for individuals with HVDs ($r_s = .74$, $p < .001$), but significant and moderate for controls ($r_s = .50$, $p = .007$; see Figure 5a). There was a significant, strong

Table 3. Univariate analysis of variance results for laryngeal kinematic trajectory measures with group (hyperfunctional voice disorder [HVD], control) as a fixed factor, following a multivariate analysis of variance that examined gesture rate, group, and their interaction.

Measure	*F*	*df*	*p*	η_p^2	Effect size interpretation	Standardized means [95% CI]	
						HVD	Control
KS	0.854	1, 155	.357	.005	Negligible	0.07 [−0.15, 0.29]	−0.08 [−0.29, 0.14]
STI	7.527	1, 155	.007	.046	Small	0.22 [−0.01, 0.45]	−0.21 [−0.42, 0.01]
AI	7.969	1, 155	.005	.049	Small	0.21 [−0.01, 0.43]	−0.22 [−0.43, 0.01]

Note. Effect size interpretations from Cohen (1988). CI = confidence interval; KS = kinematic stiffness; STI = spatiotemporal index; AI = asymmetry index.

Table 4. Back-transformed estimated marginal means for laryngeal kinematic measures based on group (hyperfunctional voice disorder [HVD], control), following a multivariate analysis of variance examining gesture rate, group, and their interaction.

Measure	Back-transformed means [95% CI]	
	HVD	Control
KS	28.69 [26.44, 31.38]	27.19 [25.27, 29.47]
STI	10.66 [9.51, 11.98]	8.58 [7.73, 9.55]
AI	2.29 [2.03, 2.59]	1.82 [1.64, 2.03]

Note. CI = confidence interval; KS = kinematic stiffness; STI = spatiotemporal index; AI = asymmetry index.

relationship between KS and AI for the group with HVDs (r_s = .63, p < .001), and a nonsignificant, weak relationship for controls (r_s = .30, p = .120; see Figure 5b). Finally, the relationship between STI and AI demonstrated a significant, very strong correlation in the group with HVDs (r_s = .87, p < .001) and significant, strong correlation for controls (r_s = .70, p < .001; see Figure 5c). For vocal effort level, the relationship between KS and STI was significant and strong for individuals with HVDs (r_s = .75, p < .001), but significant and weak for controls (r_s = .38, p = .044; see Figure 5d). Similarly, there was a significant, moderate relationship between KS and AI for the group with HVDs (r_s = .52, p = .004), and a nonsignificant, weak relationship for controls (r_s = .26, p = .183; see Figure 5e). Finally, the relationship between STI and AI demonstrated a significant, very strong correlation in both groups (HVDs: r_s = .80, p < .001; controls: r_s = .81, p < .001; see Figure 5f).

Discussion

In this study, we quantitatively investigated underlying laryngeal motor characteristics in individuals with and without HVDs using laryngeal kinematics. Results indicated all three laryngeal kinematic measures—including KS (for laryngeal muscle tension), STI (for production variability), and AI (for movement asymmetries)—were significantly larger in individuals with HVDs compared to controls across vocal effort levels. Similar trends were observed across groups when these measures were analyzed relative to gesture rate, though group differences for KS were not statistically significant. Overall, these findings provide evidence in support of our primary hypothesis that individuals with HVDs would demonstrate heightened levels of laryngeal muscle tension, production variability, and movement asymmetries compared to individuals without HVDs. However, although we anticipated that increased gesture rates or self-induced levels of vocal effort would result in changes in laryngeal motor strategies that would have effects on KS, STI, and AI, the measures did not show a statistically significant effect of vocal effort level, gesture rate, or their interactions with group. Taken together, findings suggest that individuals with HVDs use differential underlying laryngeal motor control mechanisms compared to controls when producing similar vocal targets.

KS

Prior work validating the use of KS has speculated that higher levels of tension in underlying laryngeal musculature contribute to more laryngeal stiffness in typical speakers (A. Cooke et al., 1997). Findings from the current study are in line with this investigation, as KS values were generally larger in individuals with HVDs compared to controls; these results provide further quantitative evidence that individuals with HVDs exhibit higher overall levels of laryngeal muscle tension.

Figure 3. Estimated marginal means of nontransformed data for the multivariate analysis of variance examining group (hyperfunctional voice disorder [HVD] vs. control) and gesture rate. Variables include kinematic stiffness (KS), spatiotemporal index (STI), and asymmetry index (AI). Error bars represent 95% confidence intervals. *p < .05 based on statistical analysis of the transformed data set.

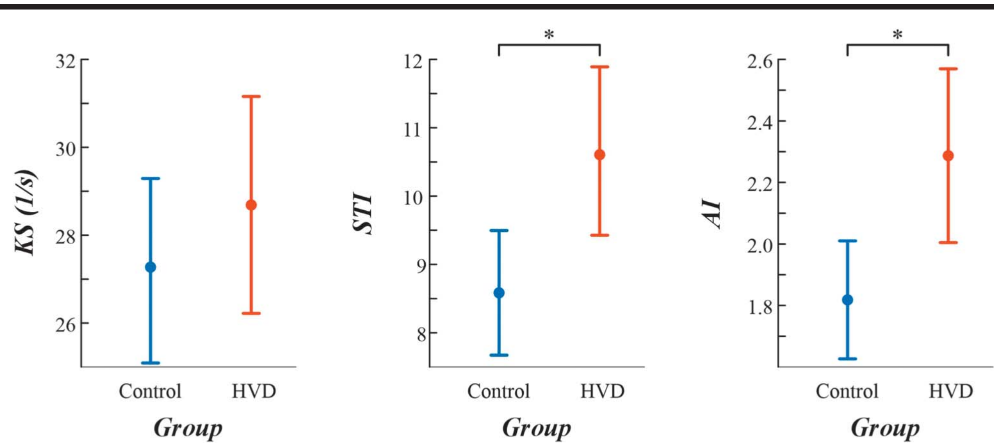


Table 5. Multivariate analysis of variance results for laryngeal kinematic trajectory measures based on group, vocal effort level, and their interaction.

Effect	Wilks' λ	*F*	*df*	*p*	η_p^2	Effect size interpretation
Intercept	.999	0.042	3, 155	.988	.001	Negligible
Group	.822	11.194	3, 155	< .001	.178	Large
Vocal effort level	.989	0.277	6, 310	.948	.005	Negligible
Group × Vocal Effort Level	.979	0.74	6, 310	.618	.014	Small

Note. Effect size interpretations from Cohen (1988).

Although no previous studies have examined average KS values directly in individuals with and without HVDs, a recent study did compare the constituents of KS in relevant groups. Crocker et al. (2024) examined 20 control speakers and two groups of individuals with HVDs (20 with muscle tension dysphonia and 20 with vocal fold nodules). Under 30-fps flexible endoscopy, each speaker produced six sustained /i/ vowels. The glottal angle waveforms were then analyzed to separately examine the three variables that compose KS (see Equation 1): θ_{max}, θ_{min}, and v_{peak}. The authors did not report any statistically significant differences between groups for any measure independently. However, they did note a trend in which values of v_{peak} were higher, but also more variable, in both HVD groups. Likewise, although it is not noted in the text of Crocker et al., visualization of their θ_{max} and θ_{min} values suggests that the difference between the two values (i.e., the movement extent) also shows a trend for lower values in the two HVD groups relative to their control group. Together, these trends for higher maximum velocities and lower movement extents suggest higher KS values in the two HVD groups. Overall, these studies indicate that the use of KS may act as a gestalt to more comprehensively characterize differences in vocal fold adduction between individuals with and without HVDs.

We did not find strong evidence to support the findings of Stepp et al. (2010) that heightened laryngeal muscle tension at baseline may limit the ability of individuals with HVDs to further increase tension. Specifically, Stepp et al. (2010) observed relatively smaller changes in KS value across gesture rates (medium to fast) in individuals with HVDs compared to controls. We speculate that the discrepancy in findings may be due to differences in the rates used in the prior study (medium = 72 BPM, fast = 104 BPM) compared to the current study (slow = 50 BPM, medium = 65 BPM, fast = 80 BPM). As our fastest rate considered was similar to medium rate in the previous study, we further speculate that a ceiling effect for KS was not approached in the current study. We chose these lower rates to reduce errors in production of participants, thus limiting the time under endoscopy, but faster rates should be explored in the future. It should also be noted that the glottal angle trajectories in Stepp et al. (2010) were generated based on sigmoidal fits of glottal angle trajectories from data collected at lower frame rates (30 fps) instead of the raw glottal angle trajectories from data collected at a sufficient frame rate to capture vocal fold posturing mechanisms (e.g., 1,000 fps as in the current study). These findings collectively suggest that the mechanisms used to self-induce vocal effort may not be comparable to real-life situational increases in vocal effort.

STI

According to motor equivalence theories and prior modeling work (Weerathunge et al., 2022), multiple kinematic configurations can produce the same acoustic output, suggesting it is crucial to characterize laryngeal kinematics to understand vocal motor variability in HVDs. In the present study, we have thus explored production variability in a more nuanced way, looking at both spatial and temporal domains in laryngeal kinematics through

Table 6. Univariate analysis of variance results for laryngeal kinematic trajectory measures with group (hyperfunctional voice disorder [HVD]) as a fixed factor, following a multivariate analysis of variance that examined vocal effort level, group, and their interaction

Measure	*F*	*df*	*p*	η_p^2	Effect size interpretation	Standardized means [95% CI]	
						HVD	Control
KS	19.896	1, 157	< .001	.112	Large	0.32 [0.11, 0.52]	−0.34 [−0.54, −0.14]
STI	30.451	1, 157	< .001	.162	Large	0.41 [0.21, 0.62]	−0.39 [−0.59, −0.19]
AI	14.661	1, 157	< .001	.085	Large	0.30 [0.09, 0.51]	−0.27 [−0.47, −0.07]

Note. Effect size interpretations from Cohen (1988). CI = confidence interval; KS = kinematic stiffness; STI = spatiotemporal index; AI = asymmetry index.

Table 7. Back-transformed estimated marginal means for laryngeal kinematic measures based on group (hyperfunctional voice disorder [HVD], control), following a multivariate analysis of variance examining vocal effort level, group, and their interaction.

Measure	Back-transformed means [95% CI]	
	HVD	**Control**
KS	32.05 [29.39, 35.25]	24.98 [23.40, 26.79]
STI	11.81 [10.62, 13.16]	7.87 [7.14, 8.69]
AI	2.44 [2.17, 2.76]	1.80 [1.62, 2.00]

Note. CI = confidence interval; KS = kinematic stiffness; STI = spatiotemporal index; AI = asymmetry index.

STI. As STI is typically used in articulatory kinematics domain (Smith et al., 2000), this is the first study to apply the STI to laryngeal kinematics.

The finding of statistically significantly higher STI values in the group with HVDs compared to controls indicates higher production variability. Our results are in line with those of previous studies that examined production variability in terms of phonatory air flow (Belsky et al., 2021; Gillespie et al., 2013; Higgins et al., 1999), voice onset time (McKenna et al., 2020), vocal intensity (Belsky et al., 2021), and tongue kinematics (Shipurkar et al., 2023). Prior altered auditory feedback studies also provide evidence that sensorimotor integration is atypical in individuals with HVDs (Abur et al., 2021), which is compatible with unstable productions compared to controls. Movement patterns driven by stable (invariant) movement plans are thought to have lower STI values, whereas movements that include real-time feedback corrections are thought to have higher STI values. The current results therefore confirm that the group with HVDs have less consistent productions and may have issues related to the stability of underlying motor commands of the target produced.

AI

This is the first application of the AI to laryngeal kinematics, which identified that individuals with HVDs produce movement trajectories that are statistically more asymmetric than do control speakers. Across all participants, AI values were typically greater than 1, suggesting that participants spent a relatively longer duration in the deceleration phase of the adductory gesture than in the acceleration phase. Since AI values were significantly higher in individuals with HVD, our findings further indicate that individuals with HVDs spend an even larger proportion of time decelerating during the adductory gesture compared to controls. More time in the deceleration phase suggests more feedback-based motor corrections in speech production (Kim & Max, 2014). If the movement is purely predefined (i.e., driven only by feedforward control), the velocity profile is hypothesized to be most efficient and thus symmetric, with AI values closer to 1. However, when the predefined movement command is not robust, feedback-based error corrections must be made to correct the ongoing production; this is observed as increased asymmetry in the velocity profile. Specifically, for closing gestures (e.g., vocal fold approximation), if the predefined motor command was insufficient, more deceleration-based changes would be needed in real time to correct the motion trajectory and stop the movement at the desired final position for voicing. This is consistent with prior research that has shown impaired updating of feedforward control in some individuals with HVDs (Abur et al., 2021). The nature of these feedback-based adjustments

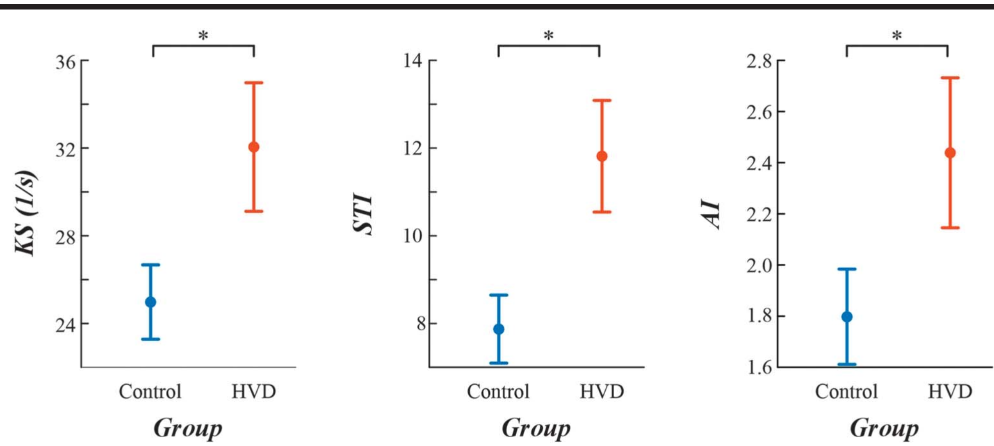
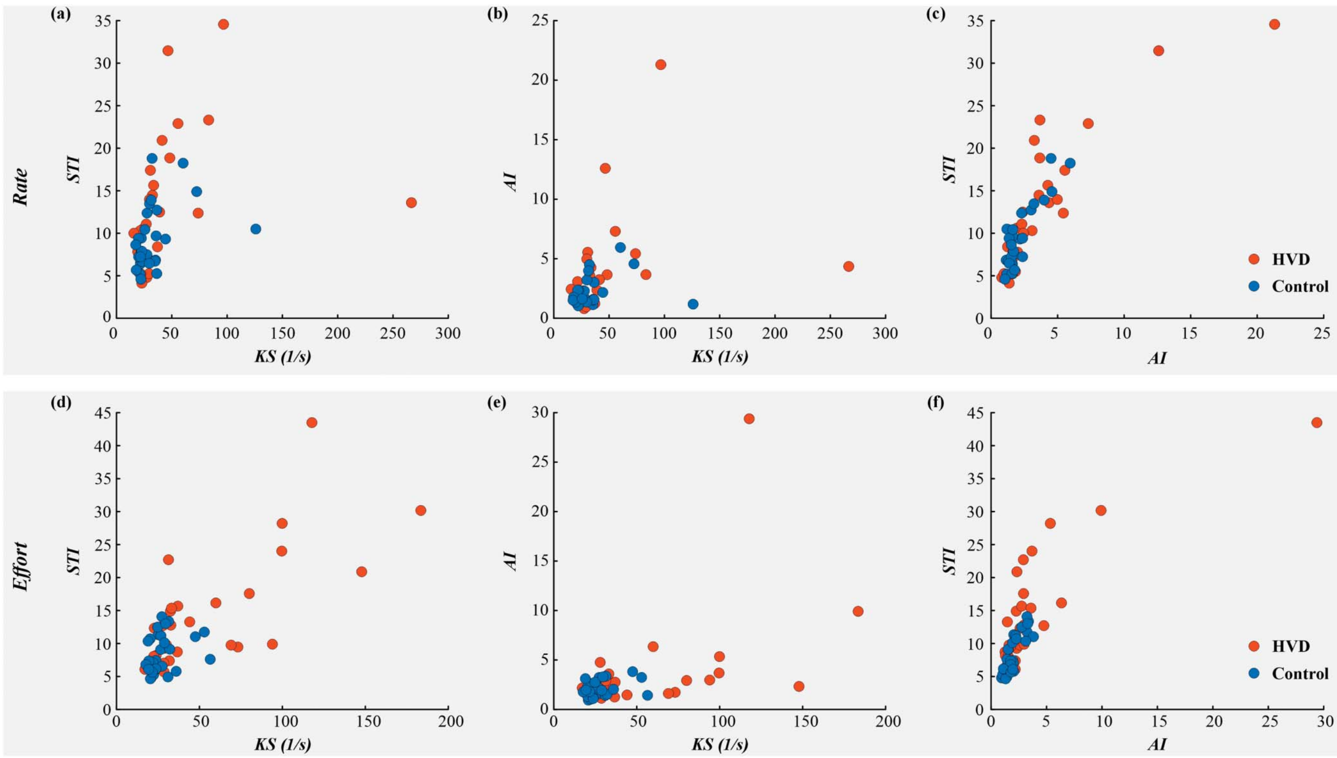


Figure 4. Estimated marginal means of nontransformed data for the multivariate analysis of variance examining group (hyperfunctional voice disorder [HVD] vs. control) and vocal effort level. Variables include kinematic stiffness (KS), spatiotemporal index (STI), and asymmetry index (AI). Error bars represent 95% confidence intervals. *$p < .05$ based on statistical analysis of the transformed data set.

Figure 5. Relationships between laryngeal kinematic measures of kinematic stiffness (KS), spatiotemporal index (STI), and asymmetry index (AI) for hyperfunctional voice disorder (HVD; orange) and control (blue) groups across (a–c) gesture rate and (d–f) vocal effort level. Each dot corresponds to a specific kinematic measure averaged across all /ifi/ instances for a given participant.



during deceleration are unlikely to be based on auditory feedback. First, although adaptive responses to altered auditory feedback are atypical in speakers with HVDs, their responses to brief, unexpected perturbations in auditory feedback thought to reflect auditory feedback control are typical (Abur et al., 2021), not larger. Second, adduction occurs in the absence of phonation, which precludes the use of auditory feedback for adjustments. Thus, feedback adjustments related to elevated values of AI are more likely to be related to somatosensory feedback. Future work is necessary to test this hypothesis using more direct perturbations of somatosensation in individuals with HVDs.

Based on prior studies on jaw, lower lip, and tongue tip elevating gestures to close the mouth (Adams, 1990; Connor et al., 1988), we expected the time spent in deceleration phase to increase as gesture rate became slower (i.e., higher AI for slower rates) and vice versa. However, we did not find evidence in support of this hypothesis, as gesture rate nor the interaction of Group × Gesture Rate were significant in the model for AI. It is possible that the duration of laryngeal adductory movements is not particularly sensitive to gesture rate modifications in the current study, perhaps due to the differing roles of auditory feedback (with its much longer delays) in laryngeal adductory movements versus articulatory movements during voiced speech (Adams, 1990).

Relationships Among Laryngeal Kinematic Measures

In an effort to understand whether the measures of laryngeal kinematic trajectories assess similar information of the underlying laryngeal motor control mechanism, we investigated the correlational relationships between them via post hoc Spearman's rank correlation coefficients. We observed that the three measures (KS, STI, AI) demonstrated significant, moderate-to-strong relationships in individuals with HVDs; however, only STI and AI exhibited a significant, very strong relationship in individuals with and without HVDs. STI characterizes production variability, whereas AI characterizes movement symmetry; although their calculations are unique, both can be interpreted in terms of feedback versus feedforward control mechanisms. Thus, the strong positive correlation between these two metrics is consistent with the interpretation that individuals without robust feedforward control are likely to be more variable across repeated productions (higher STI) and that movement duration—when considering within-speaker productions—would be biased toward the deceleration phase (higher AI). This strong correlation, in combination with the group effects in both measures, provides compelling evidence for impairments in feedforward laryngeal control in individuals with HVDs.

Limitations and Future Directions

There are several limitations in this study. First, the characteristic supraglottic constriction observed in individuals with HVDs (Roy, 2008) resulted in a comparably higher number of trials that were discarded in the group with HVDs compared to the control group. Although the difference in trial numbers were not significant, we implemented a bias correction for STI based on the number of trials considered for each calculation as a result (Wisler et al., 2022). In addition to differences in trial numbers, the sample size included in this study was not sufficient to analyze different phenotypes or pathophysiological models of HVDs; future work should replicate the current study in a larger data set with sufficient sample sizes in each phenotype (phonotraumatic and nonphonotraumatic) to characterize possible deviations in laryngeal motor control.

The CAPE-V ratings for control and HVD groups showed substantial overlap. Although this could be seen as a limitation, we opted to include participants with milder severity of dysphonia to reflect the heterogeneity of the clinical population. Diagnosis in this study considered multiple factors beyond these perceptual ratings, including patient symptoms and laryngeal examination. As this overlap may impact the results, we suggest that future studies in larger sample sizes explore the relationship between these laryngeal kinematic measures and metrics of severity.

Conclusions

We investigated three laryngeal kinematic measures to determine if there were differential effects between individuals with and without HVDs. High-speed videoendoscopic techniques were used to extract laryngeal kinematic measures of KS, STI, and AI across a range of speaking conditions that spanned different gesture rates and levels of self-induced vocal effort. Results indicated that individuals with HVDs generally exhibited higher laryngeal kinematic measures compared to controls, indicating higher laryngeal muscle tension (via KS), production variability (via STI), and movement asymmetry (via AI). These findings suggest that individuals with HVDs use different underlying laryngeal motor control mechanisms compared to vocally healthy controls when producing the same vocal targets.

Data Availability Statement

The data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This research was supported by Grants P50 DC015446 (Robert E. Hillman) and R01 DC015570 (Cara E. Stepp) from the National Institute on Deafness and Other Communication Disorders. It was also supported by a Graduate Fellow Award from the Rafik B. Hariri Institute for Computing and Computational Science and Engineering (Hasini R. Weerathunge) and an ASHFoundation New Century Doctoral Scholarship (Hasini R. Weerathunge).

References

Abur, D., Perkell, J. S., & Stepp, C. E. (2022). Impact of vocal effort on respiratory and articulatory kinematics. *Journal of Speech, Language, and Hearing Research, 65*(1), 5–21. https://doi.org/10.1044/2021_JSLHR-21-00323

Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L. F., Noordzij, J. P., & Stepp, C. E. (2021). Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders. *Scientific Reports, 11*(1), Article 13123. https://doi.org/10.1038/s41598-021-92250-8

Adams, S. G. (1990). *Rate and clarity of speech: An X-ray microbeam study* (Vol. 2). University of Wisconsin–Madison.

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research, 36*(1), 41–54. https://doi.org/10.1044/jshr.3601.41

Altman, K. W., Atkinson, C., & Lazarus, C. (2005). Current and emerging concepts in muscle tension dysphonia: A 30-month review. *Journal of Voice, 19*(2), 261–267. https://doi.org/10.1016/j.jvoice.2004.03.007

American Speech-Language-Hearing Association. (2002). *ConsensusAuditory–Perceptual Evaluation of Voice (CAPE-V) form.* https://www.asha.org/form/cape-v/

Baker, J., & Ben-Tovim, D. I. (1996). Psychological aspects of functional voice disorders. *British Journal of Disorders of Communication, 31*(3), 275–290.

Beggs, W. D. A., & Howarth, C. I. (1972). The movement of the hand towards a target. *Quarterly Journal of Experimental Psychology, 24*(4), 448–453. https://doi.org/10.1080/14640747208400304

Belsky, M. A., Rothenberger, S. D., Gillespie, A. I., & Gartner-Schmidt, J. L. (2021). Do phonatory aerodynamic and acoustic measures in connected speech differ between vocally healthy adults and patients diagnosed with muscle tension dysphonia? *Journal of Voice, 35*(4), 663.e1–663.e7. https://doi.org/10.1016/j.jvoice.2019.12.019

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

Bullock, D., & Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review, 95*(1), 49–90. https://doi.org/10.1037/0033-295x.95.1.49

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Connor, N. P., Abbs, J. H., Nychis-Florence, C., & Caligiuri, M. P. (1988). Alterations in lip and jaw kinematics as a function of speech rate. *The Journal of the Acoustical Society of America, 83*(S1), S112–S113. https://doi.org/10.1121/1.2025151

Cooke, A., Ludlow, C. L., Hallett, N., & Selbie, W. S. (1997). Characteristics of vocal fold adduction related to voice onset. *Journal of Voice, 11*(1), 12–22. https://doi.org/10.1016/s0892-1997(97)80019-7

Cooke, J. (1980). 11 The organization of simple, skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Advances in psychology: Vol. 1. Tutorials in motor behavior* (pp. 199–212). Elsevier. https://doi.org/10.1016/S0166-4115(08)61946-9

Crocker, C., Toles, L. E., Morrison, R. A., & Shembel, A. C. (2024). Relationships between vocal fold adduction patterns, vocal acoustic quality, and vocal effort in individuals with and without hyperfunctional voice disorders. *Journal of Voice*. Advance online publication. https://doi.org/10.1016/j.jvoice.2023.12.012

Dailey, S. H., Kobler, J. B., Hillman, R. E., Tangrom, K., Thananart, E., Mauri, M., & Zeitels, S. M. (2005). Endoscopic measurement of vocal fold movement during adduction and abduction. *The Laryngoscope, 115*(1), 178–183. https://doi.org/10.1097/01.mlg.0000150701.46377.df

Diaz-Cadiz, M. E., McKenna, V. S., Vojtech, J. M., & Stepp, C. E. (2018). Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy. *Journal of Speech, Language, and Hearing Research, 62*(6), 1685–1706. https://doi.org/10.1044/2019_JSLHR-S-18-0405

Dietrich, M., Verdolini Abbott, K., Gartner-Schmidt, J., & Rosen, C. A. (2008). Psychosocial aspects of voice disorders: A review of the literature. *Journal of Voice, 22*(6), 662–671. https://doi.org/10.1016/j.jvoice.2014.02.010

Eadie, T. L., & Doyle, P. C. (2005). Classification of dysphonic voice: Acoustic and auditory–perceptual measures. *Journal of Voice, 19*(1), 1–14. https://doi.org/10.1016/j.jvoice.2004.02.002

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

Freeman, E., Woo, P., Saxman, J. H., & Murry, T. (2012). A comparison of sung and spoken phonation onset gestures using high-speed digital imaging. *Journal of Voice, 26*(2), 226–238. https://doi.org/10.1016/j.jvoice.2010.11.005

Gillespie, A. I., Gartner-Schmidt, J., Rubinstein, E. N., & Abbott, K. V. (2013). Aerodynamic profiles of women with muscle tension dysphonia/aphonia. *Journal of Speech, Language, and Hearing Research, 56*(2), 481–488. https://doi.org/10.1044/1092-4388(2012/11-0217)

Gracco, V. L. (1988). Timing factors in the coordination of speech movements. *Journal of Neuroscience, 8*(12), 4628–4639. https://doi.org/10.1523/JNEUROSCI.08-12-04628.1988

Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research, 65*(1), 156–166. https://doi.org/10.1007/BF00243838

Higgins, M. B., Chait, D. H., & Schulte, L. (1999). Phonatory air flow characteristics of adductor spasmodic dysphonia and muscle tension dysphonia. *Journal of Speech, Language, and Hearing Research, 42*(1), 101–111. https://doi.org/10.1044/jslhr.4201.101

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

Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zanartu, 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

Ikuma, T., Kunduk, M., & McWhorter, A. J. (2013). Preprocessing techniques for high-speed videoendoscopy analysis. *Journal of Voice, 27*(4), 500–505. https://doi.org/10.1016/j.jvoice.2013.01.014

Iwahashi, T., Ogawa, M., Hosokawa, K., Kato, C., & Inohara, H. (2016). A detailed motion analysis of the angular velocity between the vocal folds during throat clearing using high-speed digital imaging. *Journal of Voice, 30*(6), 770.e1–770.e8. https://doi.org/10.1016/j.jvoice.2015.11.004

Kim, K. S., & Max, L. (2014). Estimating feedforward vs. feedback control of speech production through kinematic analyses of unperturbed articulatory movements. *Frontiers in Human Neuroscience, 8,* Article 911. https://doi.org/10.3389/fnhum.2014.00911

McKenna, V. S., Heller Murray, E. S., Lien, Y. S., & Stepp, C. E. (2016). The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *Journal of Speech, Language, and Hearing Research, 59*(6), 1283–1294. https://doi.org/10.1044/2016_JSLHR-S-15-0406

McKenna, V. S., Hylkema, J. A., Tardif, M. C., & Stepp, C. E. (2020). Voice onset time in individuals with hyperfunctional voice disorders: Evidence for disordered vocal motor control. *Journal of Speech, Language, and Hearing Research, 63*(2), 405–420. https://doi.org/10.1044/2019_JSLHR-19-00135

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

Mi, Q., Li, N. Y., Ziraldo, C., Ghuma, A., Mikheev, M., Squires, R., Okonkwo, D. O., Verdolini-Abbott, K., Constantine, G., An, G., & Vodovotz, Y. (2010). Translational systems biology of inflammation: Potential applications to personalized medicine. *Personalized Medicine, 7*(5), 549–559. https://doi.org/10.2217/pme.10.45

Munhall, K. G. (1984). *Temporal adjustment in speech motor control: Evidence for laryngeal kinematics*. McGill University.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance, 11*(4), 457–474. https://doi.org/10.1037//0096-1523.11.4.457

Nguyen, D. D., Chacon, A. M., Novakovic, D., Hodges, N. J., Carding, P. N., & Madill, C. (2022). Pitch discrimination testing in patients with a voice disorder. *Journal of Clinical Medicine, 11*(3), Article 584. https://doi.org/10.3390/jcm11030584

Oates, J., & Winkworth, A. (2008). Current knowledge, controversies and future directions in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology, 10*(4), 267–277. https://doi.org/10.1080/17549500802140153

Ostry, D. J., Cooke, J. D., & Munhall, K. G. (1987). Velocity curves of human arm and speech movements. *Experimental Brain Research, 68*(1), 37–46. https://doi.org/10.1007/BF00255232

Rosenthal, A. L., Lowell, S. Y., & Colton, R. H. (2014). Aerodynamic and acoustic features of vocal effort. *Journal of Voice, 28*(2), 144–153. https://doi.org/10.1016/j.jvoice.2013.09.007

Roy, N. (2008). Assessment and treatment of musculoskeletal tension in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology, 10*(4), 195–209. https://doi.org/10.1080/17549500701885577

Roy, N., Bless, D. M., & Heisey, D. (2000). Personality and voice disorders: A superfactor trait analysis. *Journal of Speech, Language, and Hearing Research, 43*(3), 749–768. https://doi.org/10.1044/jslhr.4303.749

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research, 47*(2), 281–293. https://doi.org/10.1044/1092-4388(2004/023)

Shiller, D. M., Laboissiere, R., & Ostry, D. J. (2002). Relationship between jaw stiffness and kinematic variability in speech. *Journal of Neurophysiology, 88*(5), 2329–2340. https://doi.org/10.1152/jn.00286.2002

Shipurkar, R., White, M., & Katz, W. (2023). Normalization of speech kinematic data for characterizing primary muscle tension dysphonia. *The Journal of the Acoustical Society of America, 154*(Suppl. 4), A337–A337. https://doi.org/10.1121/10.0023721

Smith, A., Johnson, M., McGillem, C., & Goffman, L. (2000). On the assessment of stability and patterning of speech movements. *Journal of Speech, Language, and Hearing Research, 43*(1), 277–286. https://doi.org/10.1044/jslhr.4301.277

Stepp, C. E., Hillman, R. E., & Heaton, J. T. (2010). A virtual trajectory model predicts differences in vocal fold kinematics in individuals with vocal hyperfunction. *The Journal of the Acoustical Society of America, 127*(5), 3166–3176. https://doi.org/10.1121/1.3365257

Stepp, C. E., Lester-Smith, R. A., Abur, D., Daliri, A., Pieter Noordzij, J., & Lupiani, A. A. (2017, June 10). Evidence for auditory-motor impairment in individuals with hyperfunctional voice disorders. *Journal of Speech, Language, and Hearing Research, 60*(6), 1545–1550. https://doi.org/10.1044/2017_JSLHR-S-16-0282

Stepp, C. E., Sawin, D. E., & Eadie, T. L. (2012). The relationship between perception of vocal effort and relative fundamental frequency during voicing offset and onset. *Journal of Speech, Language, and Hearing Research, 55*(6), 1887–1896. https://doi.org/10.1044/1092-4388(2012/11-0294)

Tam, K. H., Carding, P., Heard, R., & Madill, C. J. (2018). *The relationship between voice quality and pitch discrimination ability in a population with features of mild vocal hyperfunction* [Paper presentation]. Voice Foundation's 47th Annual Symposium, Philadelphia, PA, United States.

Titze, I. R. (1991). A model for neurologic sources of aperiodicity in vocal fold vibration. *Journal of Speech and Hearing Research, 34*(3), 460–472. https://doi.org/10.1044/jshr.3403.460

Van Houtte, E., Van Lierde, K., & Claeys, S. (2011). Pathophysiology and treatment of muscle tension dysphonia: A review of the current knowledge. *Journal of Voice, 25*(2), 202–207. https://doi.org/10.1016/j.jvoice.2009.10.009

Vojtech, J. M., & Stepp, C. E. (2024). Effects of age and Parkinson's disease on the relationship between vocal fold abductory kinematics and relative fundamental frequency. *Journal of Voice, 38*(5), 1008–1022. https://doi.org/10.1016/j.jvoice.2022.03.007

Weerathunge, H. R., Alzamendi, G. A., Cler, G. J., Guenther, F. H., Stepp, C. E., & Zanartu, M. (2022). LaDIVA: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology, 18*(6), Article e1010159. https://doi.org/10.1371/journal.pcbi.1010159

Wisler, A., Goffman, L., Zhang, L., & Wang, J. (2022). Influences of methodological decisions on assessing the spatiotemporal stability of speech movement sequences. *Journal of Speech, Language, and Hearing Research, 65*(2), 538–554. https://doi.org/10.1044/2021_JSLHR-21-00298

Ziethe, A., Petermann, S., Hoppe, U., Greiner, N., Bruning, M., Bohr, C., & Dollinger, M. (2019). Control of fundamental frequency in dysphonic patients during phonation and speech. *Journal of Voice, 33*(6), 851–859. https://doi.org/10.1016/j.jvoice.2018.07.001