Quantifying Vocal Fold Vibration in the Coronal Plane Using Optical Coherence Tomography in Normal Human Subjects[1]

[*]Diego E. Razura, [*,†]Anna M. Wisniowiecki, [‡]Zhaoyan Zhang, [*]Marcela A. Moran, [*,§,¶]Brian E. Applegate, and [*]Michael M. Johns III, *§¶*Los Angeles, ‡California, and †College Station, Texas*

SUMMARY: Objective. Assessment of human vocal fold (VF) morphology and vibration is largely limited to surface visualization via laryngeal videostroboscopy. Using optical coherence tomography (OCT), this study aims to measure VF vibratory dynamics in the coronal plane of healthy volunteers to support a quantitative assessment of VF biomechanics.

Study Design. Cross-sectional observational pilot study

Methods. Twelve healthy participants underwent awake, transoral imaging using a custom-built, hand-held rigid laryngoscope equipped with OCT technology. Phonation was captured during sustained vowel /i/ production at a modal pitch and loudness. A novel algorithm was utilized to reconstruct the phase-resolved VF vibratory cycle during phonation from the OCT data. Seven dynamic metrics were extracted from each video reconstruction, including closed quotient (CQ), angle at closure (AaC), divergent phase vertical thickness (DPVT), closed phase vertical span (CPVS), mucosal peak, amplitude of lateral excursion (amplitude), and vertical phase difference (VPD). Data were grouped by gender, and applicable metrics were stratified into right vocal fold and left vocal fold regions. Independent t-tests were performed to assess gender and laterality differences for each metric. Pairwise Bonferroni-adjusted Pearson correlations were performed between metrics.

Results. Seven novel metrics were successfully extracted and measured from OCT data in all subjects. Significant gender differences were found for DPVT ($p = 0.015$), amplitude ($p = 0.049$), and VPD ($p = 0.043$). No significant laterality differences were observed except for DPVT in the female group. Correlation analyses showed a positive correlation between DPVT and CPVS ($r = +0.89$). CQ was negatively correlated with AaC ($r = -0.74$) and positively correlated with VPD ($r = +0.83$)

Conclusion. OCT imaging of the VFs can be successfully performed in normal subjects during phonation to visualize sub-surface VF anatomy and vibration, enabling quantitative characterization of VF dynamics.

Keywords: Optical coherence tomography–Vocal fold vibration–Voice assessment–Laryngoscopy–Gender–Voice biomechanics.

INTRODUCTION

Dysphonia impacts approximately 8% of individuals in the United States at any given time and up to 30% of people throughout their lifetime, yet vocal fold (VF) clinical imaging only captures the surface of the tissue.[1,2] The VF's intricate, multi-layered mucosa and ligament generate a three-dimensional (3-D) mucosal wave essential for phonation—a complexity missed by gold-standard videostroboscopy surface-only visualization.[3–5] Additionally, videostroboscopy interpretation is a subjective process and often results in variable inter-rater reliability.[6,7] This underscores the clinical need for a non-invasive clinical tool capable of providing high-resolution imaging of the subsurface VF and vibratory biomechanics.[8,9]

Conventional imaging modalities, such as magnetic resonance imaging and computed tomography, are unable to acquire the necessary spatial and temporal resolution of the VF mucosal layers in real time during phonation.[10,11] Newer techniques such as high-speed digital imaging (HSDI), videokymography, high-speed laser projection, and ultrasonography have been explored for assessing vocal function and 3-D morphology; however, these methods are confined to the surface-level and limited by spatial uncertainty, inherent complexity of data acquisition, or reliance on subjective interpretations.[12–14] Optical coherence tomography (OCT), by contrast, can acquire depth-resolved cross-sectional images of the VFs noninvasively.[15,16] Its use continues to grow in other specialties, but the clinical feasibility and utility of OCT in laryngology practice for the functional assessment of the VFs have yet to be effectively demonstrated.[17] Prior OCT prototypes have demonstrated promise in VF imaging in awake patients, but faced challenges such as gag provocation, motion artifacts, and low temporal resolution due to frame rates below the voice fundamental frequency (90–260 Hz).[18–21]

To address these issues, we developed a 70° transoral OCT probe with a >1 kHz frame rate and simultaneous coaxial video-endoscopy. In this study, we investigate VF

Accepted for publication February 24, 2026.

[1] This article was presented at The Voice Foundation Annual Symposium: Care of the Professional Voice 54th Symposium, May 28th – June 1st, 2025.

From the *Caruso Department of Otolaryngology-Head & Neck Surgery, Keck School of Medicine of the University of Southern California, Los Angeles, California; †Department of Biomedical Engineering, Texas A&M University, College Station, Texas; ‡Department of Head and Neck Surgery, David Geffen School of Medicine at the University of California, Los Angeles, California; §Alfred Mann Department of Biomedical Engineering, University of Southern California, Los Angeles, California; and the ¶Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California.

Address correspondence and reprint requests to Michael M. Johns III, USC Department of Otolaryngology, Head and Neck Surgery, Keck School of Medicine of USC, 1537 Norfolk Street, Suite 5800, Los Angeles, California 90033. E-mail: Michael.johns@med.usc.edu

Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997

https://doi.org/10.1016/j.jvoice.2026.02.046

vibration dynamics in the coronal plane using our OCT system and reconstruction method for glottal phase-resolved, cross-sectional imaging. Specifically, we propose seven novel metrics aimed at capturing medial surface and subsurface biomechanics. We also assess potential gender-based differences within the metrics to establish quantitative baselines for future clinical applications. Ultimately, we envision that coronal-plane OCT-derived images and vibratory metrics could guide more precise diagnoses, enhance treatment planning, and provide better therapeutic outcomes for VF disorders.

MATERIALS AND METHODS

Study design and participant selection

This cross-sectional observational pilot study was conducted with approval from the Institutional Review Board of the University of Southern California (HS-20–00296). Twelve healthy participants (1:1 male-to-female ratio) were recruited from a single academic medical center. All participants self-reported no known vocal fold (VF) disorders and were able to tolerate awake transoral imaging. Imaging with the OCT hand-held rigid laryngoscope confirmed normal VF appearance. Written informed consent was obtained from each participant before enrollment and after the explanation of the minimally invasive procedure.

Imaging procedure

A custom-built, hand-held rigid laryngoscope with a 70° viewing angle was used to visualize the vocal folds. This laryngoscope was integrated with a swept-source OCT system, enabling cross-sectional imaging at > 1 kHz. For each imaging session, a fellowship-trained laryngologist served as the operator (M.J. III). Before each imaging session, the tip of the laryngoscope was soaked in an anti-fog solution to maintain clear visualization. An ophthalmic chin rest was used to stabilize the head to reduce motion artifacts during data collection, and transoral access was maintained by gently depressing the tongue with the laryngoscope. All imaging took place while participants were awake and unmedicated (i.e., without topical anesthetic agents). Each participant was instructed to maintain a sniffing position on the chin rest, characterized by neck flexion and head extension to optimize visualization of the VFs. A labeled representation of a typical imaging session setup can be found in Figure 1.

Participants were instructed to sustain the vowel /i/ at modal pitch and loudness for at least six seconds, ensuring a consistent vibration pattern for subsequent analysis. A halogen light source, delivered through a fiber bundle, illuminated the hypopharynx and larynx. To accommodate the variation in VF distance among participants, we used a variable optical delay line to enable a variable working distance range of > 5 cm (i.e., working distances of 54–110 mm from the tip of the laryngoscope). Given the high prevalence of benign lesions in the middle membranous region of the VFs,[22] we targeted this region for each imaging session. OCT data from other regions along the VFs were excluded from this study. All patient data were de-identified.

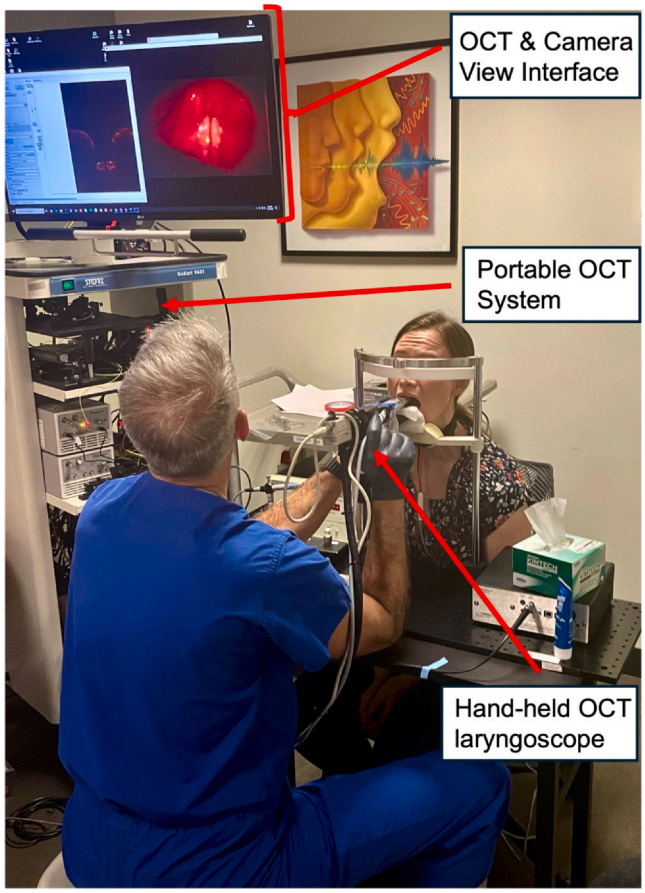

FIGURE 1. Awake transoral optical coherence tomography (OCT) setup for vocal fold imaging. The clinician operates a hand-held OCT laryngoscope while the subject sits upright with the chin stabilized on an ophthalmic chinrest. The portable imaging system provides simultaneous display of the OCT cross-section and the endoscopic camera field of view. A crosshair overlaid on the camera image displays the position of the OCT image center for real-time guidance.

OCT system design

The OCT system has been described in detail elsewhere.[23] Briefly, the OCT system was based on a Mach-Zehnder interferometer using a swept-source laser (Insight, Inc., λ_0 = 1.304 μm) with an A-line rate (laser sweep rate) of ~207 kHz. The OCT system provides an axial resolution of 22.6 μm and a lateral resolution of ~250 μm at a working distance of 70 mm. Using a fast-scanning micro-electro-mechanical (MEMs) mirror, we were able to achieve B-scan rates of ~1.02 kHz. A co-aligned video camera captured standard endoscopic video simultaneously with acquired OCT scans. All imaging data were securely stored for later post-processing and metric extraction.

Postprocessing

A novel algorithm, inspired by retrospective gating,[24] was implemented in MATLAB (MathWorks, 2022a) to create temporally and spatially resolved movies of the in vivo VF vibratory cycle from the acquired B-scans.[23] Using the

previously mentioned stabilization methods, we acquired high-speed OCT images during sustained phonation. OCT sampling was restricted to a single cross-sectional plane that produced time-resolved image sequences that depict the cyclic tissue motion at a defined position along the glottal axis. The OCT data for each participant were first screened to ensure (1) that the VFs remained stable within the OCT and endoscopic video, (2) that background noise and motion artifacts were minimal, and (3) that the VF edges were clearly demarcated. Three of the authors (D.R., A.W., M.J. III) independently verified these quality criteria before selecting a single dataset per participant. A subset of 20 consecutively acquired OCT B-Scans w then employed for reconstruction of the local phase-resolved glottal cycle.

Each reconstructed dataset comprised 150 frames representing one complete mucosal wave cycle. These frames were imported into ImageJ2 (v. 2.14.0) for measurement of vibratory metrics. Initially, both a midline and a tilt line were added to every frame. For clarity, the opening phase was defined as the portion of the glottal cycle during which the VFs separate following closure, whereas the closing phase was defined as the interval during which the VFs are in contact. To establish the midline, we first identified the frame corresponding to initial VF contact at the onset of the closed phase and placed a marker at that pixel coordinate. Additional markers were placed at the midpoint of the closed phase and at the frame of final contact. The three markers were used to visually approximate a line of best fit that closely intersected all markers. To account for any tilt introduced by the operator's hand orientation, a reference line was added between the right and left lateral VF edges, landmarks that remain relatively stable and unchanged throughout the vibratory cycle.[25] The tilt line served as an orientation reference to maintain alignment with the true imaging plane across all measurements.

Using these reference lines, seven vibratory metrics were measured for each subject, consisting of timing, distance and angle measurements. The definition and rationale for each metric are described in detail below. All measurements were recorded in millimeters (mm), except for the closed quotient (ratio), angle at closure (degrees), and vertical phase difference (degrees). For each measurement, careful consideration was taken to distinguish true tissue boundaries from any artifacts such as overlying secretions or fixed-pattern noise. The right and left VFs were measured independently, and data were subsequently grouped by participant gender. Angle and distance measurements were performed using ImageJ2 tools with a custom plugin macro that facilitated measurement capture and pixel-coordinate logging. The horizontal and vertical axis of the OCT image is referred to as the x axis and z-axis, respectively.

For distance and angle measurements, consistent comparison between subjects was achieved by measuring and correcting for the deviation of the OCT scan axis from a perpendicular cross-section of the vocal folds with respect to the subject glottal axis (i.e., a 90° scan angle to the glottal axis). Respective scan angles were measured using the co-registered endoscopic camera in ImageJ2 with values ranging from α=3.1–19.8°. The angular deviation for each subject was used to calculate a scaling factor, $\cos(\alpha)$, multiplied with the horizontal (x) coordinate of each measurement to project the OCT measurement into the plane perpendicular to the glottal axis.

In addition, accurate distance measurements using OCT require consideration of the medium in which depth is measured. All metrics were measured using reconstructed images scaled for the refractive index of soft tissue (n=1.4), providing a pixel dimension of 12.75 µm x 12.75 µm. However, of the selected metrics, only mucosal peak is defined as a distance measured in tissue. The remaining metrics represent distances measured in air (n=1). Therefore, the vertical (z) components of recorded coordinates for a given metric were rescaled using the refractive index defined for the metric (n=1.4 or n=1) prior to calculation of reported angle and distance values.

The reliability and consistency of each vibratory metric were validated by recording the measurements independently on two separate days for the first four participants' datasets. Once consistent results were confirmed, the remaining participants' metrics were recorded only once using the established measurement approach.

Metric: closed quotient (CQ)

The closed quotient (CQ) is defined as a ratio of the duration where the glottis remains closed (i.e., VFs are in contact with one another) divided by the total duration of one mucosal wave cycle.[26] In our post-processing, the duration is measured using cross-sectional OCT frames that depict these two instants. Although CQ is not a new parameter for VF assessment,[6] its application to coronal plane imaging is novel. Of note, CQ is sometimes inversely discussed in the form of open quotient, where a larger CQ would correspond to a smaller open quotient, and vice versa.[26] Prior research has suggested a direct link between VF thickness and higher CQ values.[27] This implication supports the use of CQ as a baseline measure of person-to-person variability that is independent of variables such as gender and age, with the understanding that CQ can significantly be altered in response to different phonatory postures. Given this association with VF thickness, we anticipated a relationship between CQ and another novel metric, the divergent phase vertical thickness (DPVT).

To calculate CQ in our datasets, we isolated the frames corresponding to the closed phase (VFs in contact). The start of the closed phase was marked as the frame number when the lower VFs established contact, and the end was taken as the frame number in which the upper folds separated laterally (i.e., the onset of opening). The total number of closed phase frames from the start and end were accounted for and divided by the total number of frames in a full vibratory cycle (150 frames in our reconstruction). The resulting ratio represents the CQ for that cycle.

$$CQ = \frac{number\ of\ closed\ phase\ frames}{number\ of\ frames\ in\ a\ full\ vibratory\ cycle}$$

Metric: angle at closure (AaC)

The angle at closure (AaC) is defined as the angle formed between the vertical glottal midline and the superior medial edge of the VFs during the most divergent phase of closure. This metric is intended to serve as a baseline reference for evaluating asymmetries, particularly in cases where VF contour irregularities might compromise complete glottal closure. Because asymmetry is mathematically dependent on AaC, we treat it as a secondary finding rather than an independent metric.

For each measurement, a single frame was selected that best represented the onset of the closing phase. This frame was identified as the frame demonstrating the greatest inferior divergence. At this phase, the angle was calculated between the reference line and a line drawn from the inferior-medial VF edge toward the medial edge of the superior VF (Figure 2A). Asymmetric values were determined by calculating the difference between the left and right VF angles and dividing the result by two to account for contributions from both sides. A perfectly symmetric configuration will yield an asymmetry value of 0, reflecting identical angles on both sides.

Metric: divergent phase vertical thickness (DPVT)

The divergent phase vertical thickness (DPVT) is defined as an indirect measurement of the VF thickness derived from the medial edge contour during the most divergent phase of closure. Unlike traditional thickness measurements obtained while the VFs are static and not phonating, this metric evaluates thickness during dynamic motion.

The frame corresponding to either the most divergent closing phase or the initial closing phase was first identified. From this frame, a line was drawn from the superior medial edge of each VF to the midline. To accommodate potential variations in the operator's positioning of the hand-held OCT device, this line was required to run parallel to the tilt line. The DPVT was determined by measuring the distance along the z-axis from the superior medial VF edge to the inferior medial VF edge immediately parallel to the midline (Figure 2B).

Metric: closed phase vertical span (CPVS)

The closed phase vertical span (CPVS) is defined as the vertical distance along the midline between the tissue surface at the end of the closed phase and that at the initial closed phase. The end of the closed phase corresponds to the frame in which the VFs remain in contact just before transitioning into the open phase.

To calculate CPVS, the initial closed phase and the end of the closed phase were first identified. A substack comprising these two frames was created and merged into an overlay. The overlay can be optimized in ImageJ2 by using a grayscale color scheme for the coronal plane reconstruction, as we intentionally did for all datasets. Within the overlay, the distance was measured between the superior medial aspect of the VF at the end of the closed phase and the inferior medial aspect of the VF at the initial closed phase. This measurement was performed closest to the midline, aligned parallel to the midline (Figure 3).

Metric: mucosal peak

The mucosal peak is defined as the maximum vertical displacement of the VF surface relative to the tilt line. Because the mucosal peak cannot be reliably visualized from a

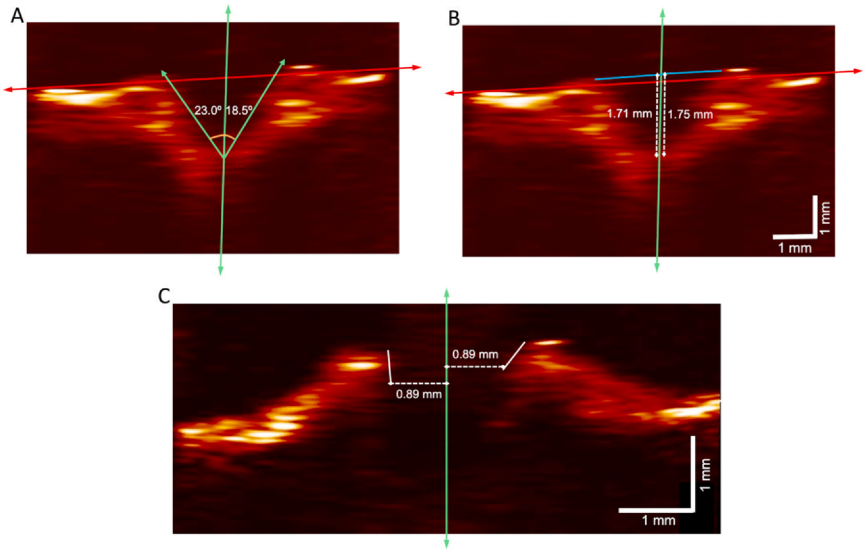

FIGURE 2. Coronal optical coherence tomography (OCT) reconstructions of the VFs with metric measurement overlays. (A) Angle at closure: the intersection of the midline (long green) and the tilt reference (red), with diverging rays (short green) drawn to each superior medial edge to measure bilateral closure angles. (B) Divergent phase vertical thickness: the perpendicular distance (dashed white) from the point of divergence to the superior medial surface (solid blue). (C) Amplitude of lateral excursion: the peak medial edge displacement (dashed white) measured 0.5 mm below the superior contour (solid white) relative to the midline.

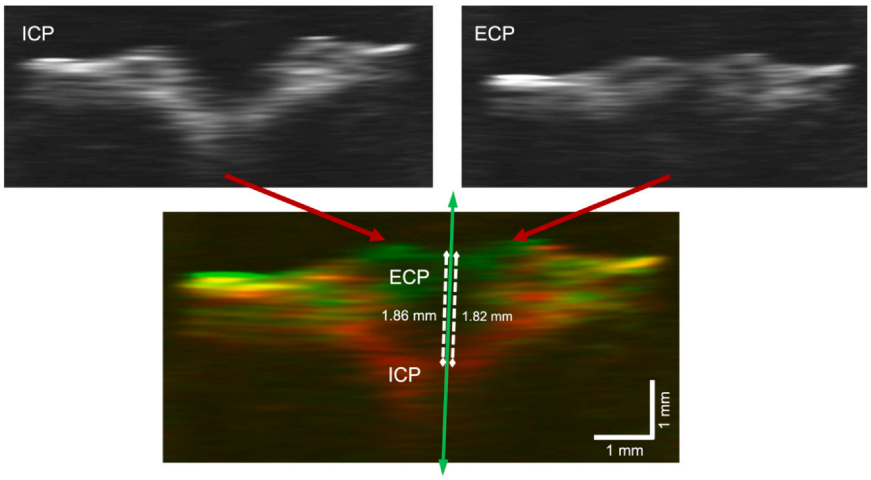

FIGURE 3. Coronal optical coherence tomography (OCT) overlay of initial closed phase (ICP) and end of closed phase (ECP) VF positions with closed phase vertical span (CPVS) measurements. The green line marks the midline. White dashed lines indicate the vertical distance between ICP and ECP margins on each VF.

superior surface view, coronal plane imaging is required. Consequently, this feature remains understudied, and its physiologic significance is poorly understood. We aim to address this previous imaging limitation with the use of OCT's cross-sectional imaging capability.

It was crucial to distinguish the true superior tissue contour from any overlying secretions, which can obscure the epithelial borders. This process involves cross-referencing preceding and subsequent frames to confirm that the traced boundary consistently followed the tissue surface rather than transient fluid layers. Furthermore, this metric is highly dependent on establishing a reliable tilt line that is derived from the stable lateral portions of the VFs, which serves as the baseline for vertical displacement measurements. To calculate the mucosal peak, the frame in which the VF surface reaches its maximum height must be identified. The measurement is then performed by determining the distance between the topmost edge of the VF surface and the tilt line. The measurement line must be orthogonal to the tilt line (Figure 4).

Metric: amplitude of lateral excursion (amplitude)

The amplitude of lateral excursion (amplitude) is defined as the lateral motion of the mucosal wave measured along the x-axis of the OCT image. Historically, amplitude has been described by the lateral motion seen from a superior view.[6,28] As shown in the literature, amplitude is directly related to vocal intensity and quality, depending on VF tissue pliability.[29] OCT may provide more precise amplitude measurements and help validate these findings.

For this metric, it is important to place extra consideration on frame selection and the depth of the amplitude measurement from the superior surface to ensure consistency. The frame chosen must have both the upper and lower vocal fold edges reach their greatest lateral distance from the glottal midline. On the chosen frame, the upper vocal fold contour is identified. An additional short 0.5 mm line in the z-axis is drawn along the medial edge. The line length was standardized across measurements to ensure consistency and capture adequate VF depth for this analysis. From the lower tip of this 0.5 mm segment, a horizontal measurement to the midline is made. This horizontal distance represents the VF's amplitude of lateral excursion (Figure 2C).

Metric: vertical phase difference (VPD)

Vertical phase difference (VPD) is the phase shift between the motion of the upper and lower margins of the vocal folds, measured in degrees of a 360° mucosal wave cycle. First described by Titze et al.,[30] VPD describes the phenomenon of time-delayed movements between the upper and lower margins of the VFs noted when one margin is open and the other is closed during phonation.[27] The phase delayed motion is important in allowing VF oscillations to be self-sustained based on exhalatory airflow energy transfer to VF tissue.[9,30,31] As it has been applied to mostly laryngeal models in vitro, the utility of OCT's cross-sectional imaging could provide the capabilities to quantify VPD clinically.

In our OCT reconstructions, we measure the VPD by identifying the two frames in which the upper and lower margins first contact each other during the closed phase. The frame index difference (Ds) is divided by the frequency corresponding to the time between reconstructed frames (fps, frames/sec). This value is then multiplied by 360° and the reconstruction frequency of the glottal cycle (f_{recon}, Hz) to acquire phase shift or VPD.

$$phase\ shift = 360° \times \left(\frac{Ds}{fps}\right) \times f_{recon}$$

Statistical analysis

All statistical analyses were performed using Stata/SE (v. 18.0, StataCorp. LLC, College Station, TX). Descriptive statistics summarized participant demographics and metric measurements. For comparison of genders, we averaged the left and

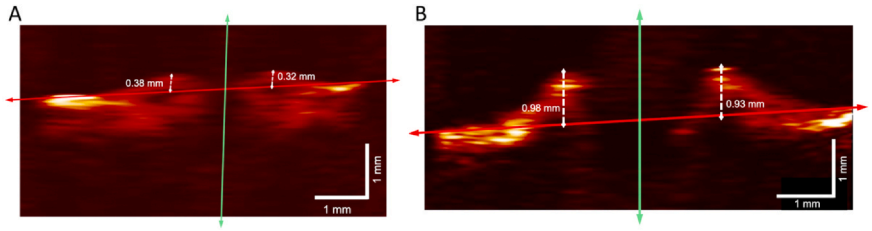

FIGURE 4. Coronal optical coherence tomography (OCT) reconstructions showing mucosal peak measurements in two subjects—(A) Volunteer 4; (B) Volunteer 5. For both panels, a glottal midline (green) and tilt reference line (red) were added. The maximum vertical mucosal peak (dashed white) was measured separately for the left and right VFs.

right vocal fold measurements for any metric that had laterality. This produced a single value per metric per subject to avoid artificial inflation of sample size. Independent two-sample t-tests (assuming equal variances) were then used to assess gender differences for all metrics. For metrics with laterality, we similarly used two-sample t-tests to identify possible differences between left and right vocal folds.

Measurement reliability was evaluated in a subset of the first four subjects by quantifying the agreement between two independent measurement attempts using both Bland-Altman and intraclass correlation (ICC) analyses.[32] For each metric, Bland-Altman analysis was performed by calculating the difference between the first and second attempts and plotting that difference as a function of the mean of the two attempts.[32,33] ICC was estimated using a one-way random-effects model to report the point estimate and its 95% confidence interval (CI) for each metric. ICC values ≥ 0.75 are considered "good" and ≥ 0.90 "excellent."[32,34] A P value of < 0.05 was considered statistically significant.

RESULTS

Twelve healthy volunteers were imaged using the OCT hand-held device to assess morphological and functional aspects of human VFs. The cohort included six male (mean age 33.7 years, range 26–53) and six female (mean age 32.5 years, range 27–38) subjects. All seven metrics were successfully measured in every subject.

Reliability

Bland-Altman analyses demonstrated good agreement between repeated measurements across all metrics, with minimal bias as reflected by small mean differences between attempts (Supplemental 1). The 95% limits of agreement were narrow relative to the magnitude of each metric, indicating that variability between repeated measurements was small and that the measurements were highly consistent. For ICC, all metrics showed excellent reliability in replicability, ranging from 0.91 to 0.99 (Supplemental 2).

Gender differences

When stratified by gender and compared using independent two-sample t-tests, three metrics demonstrated significant differences (Table 1) between genders. Male subjects had a larger DPVT ($P = 0.012$), amplitude ($P = 0.049$), and VPD ($P = 0.043$) compared with female subjects. Figure 5 shows a boxplot representation of gender differences.

Left-right vocal fold differences

The comparison of the left and right vocal folds across all volunteers and the male subgroup did not reveal any significant laterality differences for any of the metrics (Table 2). Among female subjects, we did not find laterality differences except for DPVT ($p = 0.019$).

Metric correlation analysis

Pairwise Bonferroni-adjusted Pearson correlations are summarized in Figure 6. CQ is strongly negatively correlated to AaC ($r = -0.74$) and positively correlated to VPD ($r = +0.83$). DPVT and CPVS are positively correlated ($r = +0.89$). AaC demonstrated a moderate positive correlation with amplitude ($r = +0.49$). Likewise, VPD demonstrated a moderate positive correlation with DPVT ($r = +0.54$).

DISCUSSION

Evaluation of the VFs during phonation is presently hindered by reliance on superficial two-dimensional imaging and subjective assessments. This contributes to significant inter-rater variability and limitations on our knowledge of the effects of therapeutic interventions and pathology on VF vibratory biomechanics. In this pilot study, we developed and extracted a set of hypothesis-driven vibratory metrics from coronal OCT reconstructions to characterize VF biomechanics during phonation. These metrics were motivated by the ability of OCT to resolve depth-dependent tissue motion in the coronal plane and enable quantification of vibratory features that are difficult or not possible to measure reliably with conventional endoscopic techniques. Because these metrics are novel in the context of OCT for assessing vibratory mechanics, this study also required the development and validation of reproducible measurement methods. Although not all metrics showed gender differences, the integration of these measures into VF assessment holds promise for establishing more precise person-specific baselines and enhancing our understanding of VF biomechanics, as well as providing new information toward characterizing and identifying pathologic conditions.

Multiple groups have demonstrated that OCT can visualize coronal-plane VF motion in cadaveric and in vivo settings and have adapted traditional surface-view metrics

TABLE 1.
Gender-Based Comparison of Vocal Fold Vibratory Metrics

Metric	Mean Measurement (SD)	95% Confidence Intervals	*P* value
Closed Quotient (CQ, ratio)			0.374
Male (N = 6)	0.53 (0.03)	[0.50, 0.56]	
Female (N = 6)	0.48 (0.12)	[0.35, 0.61]	
Angle at Closure (AaC, degrees)			0.551
Male (N = 6)	29.54 (4.51)	[24.80, 32.27]	
Female (N = 6)	33.43 (14.78)	[17.91, 48.94]	
Asymmetry (degrees)			0.824
Male (N = 6)	3.33 (1.88)	[1.36, 5.30]	
Female (N = 6)	3.57 (1.72)	[1.76, 5.37]	
Divergent Phase Vertical Thickness (DPVT, mm)			0.012*
Male (N = 6)	2.49 (0.43)	[2.05, 2.94]	
Female (N = 6)	1.68 (0.50)	[1.16, 2.20]	
Closed Phase Vertical Span (CPVS, mm)			0.050
Male (N = 6)	2.29 (0.65)	[1.61, 2.97]	
Female (N = 6)	1.63 (0.33)	[1.29, 1.97]	
Mucosal Peak (mm)			0.294
Male (N = 6)	0.82 (0.43)	[0.36, 1.27]	
Female (N = 6)	0.59 (0.26)	[0.32, 0.86]	
Amplitude of Lateral Excursion (Amplitude, mm)			0.049*
Male (N = 6)	0.76 (0.09)	[0.66, 0.85]	
Female (N = 6)	0.56 (0.20)	[0.35, 0.77]	
Vertical Phase Difference (VPD, degrees)			0.043*
Male (N = 6)	98.8 (10.23)	[88.06, 109.54]	
Female (N = 6)	74.8 (23.34)	[50.31, 99.29]	

**P* value < 0.05 was considered statistically significant.

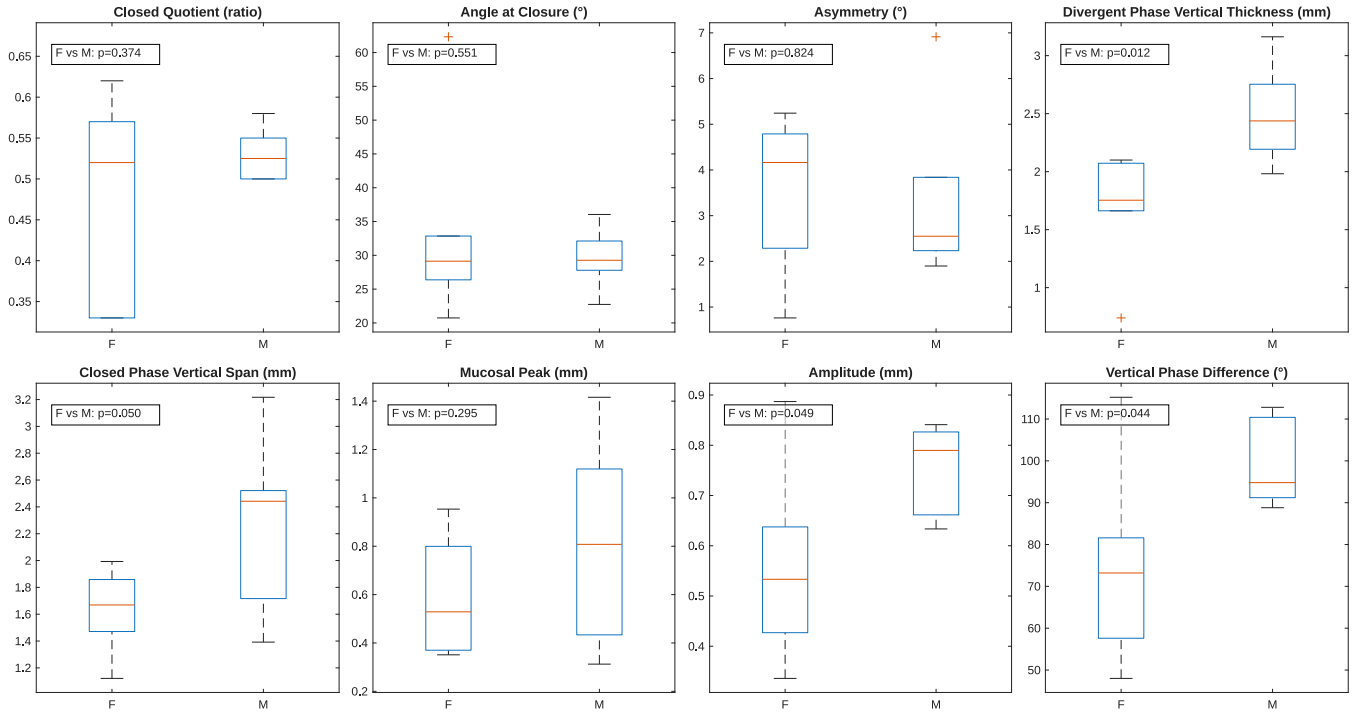

FIGURE 5. Box and whisker plots comparing female (F) and male (M) volunteers for each metric. The red line marks the median, and the box spans the interquartile range (25th-75th percentiles). Individual outliers are shown as crosses.

TABLE 2.
Left Vocal Fold (LVF) vs. Right Vocal Fold (RVF) for Metrics With Laterality

	LVF		RVF		
Metric	**Mean (SD)**	**95% CIs**	**Mean (SD)**	**95% CIs**	***P* value**
Angle at closure (AaC, degrees)					
All Volunteers (N = 12)	30.03 (9.92)	[23.73, 36.34]	32.93 (12.42)	[25.04, 40.82]	0.202
Male (N = 6)	28.36 (4.37)	[23.77, 32.95]	30.72 (7.21)	[23.15, 38.28]	0.492
Female (N = 6)	31.71 (13.81)	[17.21, 46.21]	35.14 (16.60)	[17.72, 52.57]	0.322
Divergent phase vertical thickness (DPVT, mm)					
All Volunteers (N = 12)	2.17 (0.66)	[1.75, 2.59]	2.01 (0.66)	[1.62, 2.39]	0.098
Male (N = 6)	2.59 (0.52)	[2.04, 3.14]	2.40 (0.43)	[1.94, 2.85]	0.335
Female (N = 6)	1.75 (0.49)	[1.23, 2.26]	1.61 (0.50)	[1.09, 2.14]	0.019*
Closed phase vertical span (CPVS, mm)					
All Volunteers (N = 12)	1.96 (0.61)	[1.57, 2.34]	1.96 (0.59)	[1.58, 2.33]	0.995
Male (N = 6)	2.30 (0.65)	[1.62, 2.98]	2.28 (0.65)	[1.60, 2.96]	0.484
Female (N = 6)	1.62 (0.33)	[1.27, 1.97]	1.64 (0.32)	[1.30, 1.98]	0.170
Mucosal peak (mm)					
All Volunteers (N = 12)	0.72 (0.37)	[0.49, 0.96]	0.68 (0.36)	[0.45, 0.91]	0.304
Male (N = 6)	0.85 (0.44)	[0.39, 1.32]	0.78 (0.44)	[0.32, 1.24]	0.340
Female (N = 6)	0.59 (0.25)	[0.32, 0.86]	0.59 (0.26)	[0.31, 0.86]	0.833
Amplitude of lateral excursion (amplitude, mm)					
All Volunteers (N = 12)	0.64 (0.19)	[0.52, 0.77]	0.67 (0.21)	[0.54, 0.81]	0.627
Male (N = 6)	0.75 (0.11)	[0.64, 0.87]	0.76 (0.18)	[0.57, 0.95]	0.939
Female (N = 6)	0.53 (0.21)	[0.32, 0.75]	0.58 (0.22)	[0.35, 0.82]	0.508

**P* value < 0.05 was considered statistically significant.

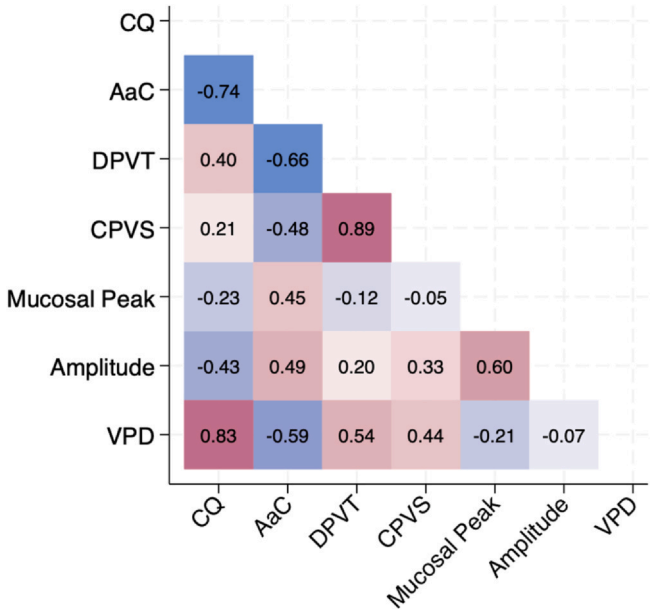

FIGURE 6. Heatmap of Bonferroni-adjusted Pearson correlations among seven vocal fold vibratory metrics: CQ (closed quotient), AaC (angle at closure), DPVT (divergent phase vertical thickness), CPVS (closed phase vertical span), mucosal peak, amplitude (amplitude of lateral excursion), and VPD (vertical phase difference). Color hue and intensity indicate correlation direction (blue = negative, red = positive) and magnitude.

to incorporate time-resolved measurements within the x-z imaging plane.[19,20,35–37] We build on this foundation by defining seven OCT metrics that directly resolve the glottal phase and are quantified using objective measurement criteria, enabling more accurate and clinically adaptable assessment of VF vibration. In an initial four-subject reliability check, all metrics exhibited negligible Bland-Altman biases with narrow limits of agreement and excellent ICCs (> 0.90). These results confirmed that our OCT metric protocols yield consistent and reproducible measurements, which justified their usage in the entire cohort. Although the initial reliability evaluation included only four subjects, the absence of significant differences across all metrics supports the metric potential for clinical implementation and further research studies.

The male cohort exhibited a significantly greater DPVT. Males are generally considered to have thicker, longer VFs that support lower fundamental frequencies.[38,39] DPVT does not directly measure static tissue thickness. Instead, it indirectly measures the dynamic vertical thickness of the VF medial surface during motion of the medial VF surface during phonation. Prior work has demonstrated that medial surface thickness has a strong influence on voice source spectral shape and voice quality.[40,41] In addition, a thicker medial surface has been linked to a greater VPD, which was noted in our male cohort.[41] However, the Bonferroni-adjusted correlation between VPD and DPVT

had only a moderate positive relationship. This could suggest that these metrics, while both increased in male participants, capture partly distinct aspects of mucosal dynamics.

CQ is theoretically positively associated with VPD, which was strongly noted in our correlation data.[41] Despite theoretical associations between VPD and CQ with male VFs, we saw no significant gender differences in the CQ metric. It is difficult to discern why CQ showed no differences. A possible reason would be inadequate resolution of the lower margin of the VFs, which would show false readings despite reliable and consistent metric methodology. Inspection of individual data revealed greater CQ variability among the female cohort, hinting that outlier values may have obscured a true mean difference. A larger cohort analysis would be needed to confirm these findings. When compared with the CQ values reported by Holmberg et al. for normal voices, the male and female groups in that study demonstrated smaller mean CQ values of 0.22 and 0.28, respectively.[26] These differences between studies may reflect methodological and task-related factors, as participants in the Holmberg et al study. were instructed to produce the syllable /pæ/, which may differ from the phonated syllable /i/ in our study. In addition, CQ in their study was derived from glottal airflow measurements, rather than from direct visualization. Future studies can explore the extent CQ changes with respect to voice conformations and vowels.

Male subjects also demonstrated greater amplitude. These findings can possibly be attributed to differences in VF length.[39] However, it has been noted that vibration amplitude is largely dependent on the region of the vocal folds, with the middle corresponding to having the largest vibration.[42] This confounding factor was addressed in our study by only processing images taken from the middle of the musculomembranous region of the VFs. Although nearly statistically significant, our male subjects had a larger CPVS compared to female subjects. In addition, we found a strong CPVS to DPVT correlation. Although CPVS is novel and uncharacterized, it may prove valuable for distinguishing across different VF conformations. Establishing whether vibratory metrics reliably differentiates male and female VFs is important because structural and functional differences contribute to characteristic pitch ranges respective of each gender.[43] In addition, these implications can be helpful in diagnosis, treatment planning, and understanding gender-related risks in voice disorders.

We did not recognize any statistically significant differences in laterality except for DPVT in the female cohort. The small asymmetries noted in our findings may be attributed to probe positioning or natural minor asymmetries commonly found in most people.[44] Given the small female sample size and the absence of consistent left and right differences, this isolated DPVT difference is more likely attributable to local variability in tissue vibration rather than a true anatomical asymmetry. Our data can support the assumption of bilateral general symmetry found in healthy VFs for most OCT-derived metrics.

Our study has several limitations. As a pilot study, the sample size was modest, limiting statistical power and restricting the generalizability of the findings. Larger and more diverse cohorts will be necessary to establish healthy baseline measurements and to strengthen cross-correlations with other imaging modality measurements. OCT imaging in vivo can be sensitive to head movement, probe positioning, and stabilization constrains, which may introduce motion artifacts. The current implementation required careful patient stabilization, manual frame selection, and manual segmentation of tissue boundaries. As such, the workflow is labor-intensive and reflects a research-stage protocol rather than an ideal clinical workflow application. Our current OCT system has limited spatial resolution, and image signal-to-noise ratio was seen to vary among subjects. For some subjects, faint medial edge features were difficult to distinguish from noise, particularly for metrics such as VPD and CQ that rely on precise boundary localization. Although objective criteria were used for frame selection and line construction, manual identification of medial edge landmarks introduces potential operator dependence. Despite these inherent procedural limitations, the strong ICCs and minimal Bland-Altman bias observed in this study suggests promise in the standardized approach to yield reproducible measurements.

Future developments will involve improving the OCT system resolution, motion correction strategies, and semi-automated or automated algorithms to reduce operator burden and enhance clinical application. Given the relative novelty of OCT within laryngology, comparisons were made to prior studies using excised laryngeal models, theoretical computer simulations, or measurements from alternative imaging modalities. Differences in methodology limit direct cross-study comparisons and underscore the need for continued validation in larger cohorts. While the study was not designed to establish immediate clinical implementation, it provides a reproducible quantitative framework for assessing depth-resolved VF vibratory mechanics. Further technological refinement and workflow efficiency will be required before routine adoption can be realized.

Conclusion

Seven novel vibratory metrics were developed for the objective measurement of VF function in the coronal plane using cycle-resolved OCT reconstructions. These proposed metrics were shown to be feasible with excellent reliability across repeated measurements, allowing characterization of normal variability, left-right VF symmetry, and relationships among depth-resolved vibratory features in healthy subjects. Within the quantitative framework, male participants demonstrated significantly higher DPVT, amplitude, and VPD. Although the clinical utility and physiological relevance of these findings require further investigation, this work established a foundation for coronal plane VF assessment. Future work should expand the diversity of the

participant pool to further evaluate the diagnostic and therapeutic potential of the OCT-derived vibratory metrics.

Funding: National Institute on Deafness and Other Communication Disorders (NIDCD) of the National Institutes of Health (NIH) (Grant No. NIH/NIDCD R25 DC019700).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

D.R. gratefully acknowledges support from the National Institute on Deafness and Other Communication Disorders (NIDCD) of the National Institutes of Health (NIH) (Grant No. NIH/NIDCD R25 DC019700).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jvoice.2026.02.046.

References

1. Bhattacharyya N. The prevalence of voice problems among adults in the United States. *Laryngoscope.* Oct 2014;124:2359–2362. https://doi.org/10.1002/lary.24740.
2. Cohen SM. Self-reported impact of dysphonia in a primary care population: an epidemiological study. *Laryngoscope.* Oct 2010;120:2022–2032. https://doi.org/10.1002/lary.21058.
3. Connor NP, Bless DM. *Videostroboscopy Textbook Laryngol.* 2007:75–84.
4. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr.* 1974;26:89–94. https://doi.org/10.1159/000263771.
5. Cohen SM, Kim J, Roy N, et al. Direct health care costs of laryngeal diseases and disorders. *Laryngoscope.* Jul 2012;122:1582–1588. https://doi.org/10.1002/lary.23189.
6. Poburka BJ, Patel RR, Bless DM. Voice-vibratory assessment with laryngeal imaging (VALI) form: reliability of rating stroboscopy and high-speed videoendoscopy. *J Voice.* 2017;31.513.e1-513.e14.
7. Rosow DE, Sulica L. Laryngoscopy of vocal fold paralysis: evaluation of consistency of clinical findings. *Laryngoscope.* Jul 2010;120:1376–1382. https://doi.org/10.1002/lary.20945.
8. Hamilton NJ. The life-cycle and restoration of the human vocal fold. *Laryngoscope Investig Otolaryngol.* 2023;8:168–176.
9. Zhang Z. Mechanics of human voice production and control. *J Acoust Soc Am.* 2016;140:2614–2635.
10. Ahmad M, Dargaud J, Morin A, Cotton F. Functional morphology of phonation evaluated by dynamic MRI. *Surg Radiologic Anatomy.* 2006;28:481–485.
11. Slonimsky E. Laryngeal imaging. *Oper Techn Otolaryngol Head Neck Surg.* 2019;30:237–242.
12. Krausert CR, Olszewski AE, Taylor LN, et al. Mucosal wave measurement and visualization techniques. *J Voice.* 2011;25:395–405.
13. Semmler M, Kniesburges S, Parchent J, et al. Endoscopic laser-based 3D imaging for functional voice diagnostics. *Appl Sci.* 2017;7:600.
14. Kandil E, Deniwar A, Noureldine SI, et al. Assessment of vocal fold function using transcutaneous laryngeal ultrasonography and flexible laryngoscopy. *JAMA Otolaryngol Head Neck Surg.* 2016;142:74–78.
15. Maguluri G, Mehta D, Kobler J, et al. Synchronized, concurrent optical coherence tomography and videostroboscopy for monitoring vocal fold morphology and kinematics. *Biomedical Optics Express.* 2019;10:4450–4461.
16. Drexler W, Liu M, Kumar A, et al. Optical coherence tomography today: speed, contrast, and multimodality. *J Biomed Opt.* 2014;19.071412-071412.
17. Tang JC, Magalhães R, Wisniowiecki A, et al. *Optical coherence tomography technology in clinical applications. Biophotonics and Biosensing*. Elsevier; 2024:285–346.
18. Baken RJ, Orlikoff RF. *Clinical measurement of speech and voice.* Cengage Learning; 2000.
19. Sharma GK, Chen LY, Chou L, et al. Surface kinematic and depth-resolved analysis of human vocal folds in vivo during phonation using optical coherence tomography. *J Biomed Opt.* Aug 2021;26(8) https://doi.org/10.1117/1.jbo.26.8.086005.
20. Coughlan CA, Chou L-d, Jing JC, et al. In vivo cross-sectional imaging of the phonating larynx using long-range Doppler optical coherence tomography. *Sci Rep.* 2016;6:22792.
21. Saggio G, Costantini G. Worldwide healthy adult voice baseline parameters: a comprehensive review. *J Voice.* 2022;36:637–649.
22. Johns MM. Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts. *Curr Opin Otolaryngol Head Neck Surg.* 2003;11:456–461.
23. Wisniowiecki A.M., Kim W., Walker C.B., et al. In vivo imaging and reconstruction of the human vocal folds during phonation with a hand-held optical coherence tomography laryngoscope. 2025.
24. Gargesha M, Jenkins MW, Wilson DL, Rollins AM. High temporal resolution OCT using image-based retrospective gating. *Opt Express.* 2009;17:10786–10799.
25. Miri AK. Mechanical characterization of vocal fold tissue: a review study. *J Voice.* 2014;28:657–667.
26. Holmberg EB, Hillman RE, Perkell JS. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am.* 1988;84:511–529.
27. Zhang Z. Vocal fold vertical thickness in human voice production and control: a review. *J Voice.* Sep 2025;39:1183–1191. https://doi.org/10.1016/j.jvoice.2023.02.021.
28. Gelfer MP, Bultemeyer DK. Evaluation of vocal fold vibratory patterns in normal voices. *J Voice.* 1990;4:335–345.
29. Döllinger M, Gómez P, Patel RR, et al. Biomechanical simulation of vocal fold dynamics in adults based on laryngeal high-speed videoendoscopy. *PLoS One.* 2017;12(11):e0187486.
30. Titze IR, Jiang JJ, Hsiao T-Y. Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. *Ann Otol Rhinol Laryngol.* 1993;102:58–63.
31. Švec JG, Kumar SP, Vencálek O, et al. Vocal fold kinematics and convergent–divergent oscillatory glottis: basic insights using mucosal wave modeling and synthetic kymograms. *J Speech Language Hearing Res.* 2025;68:1602–1617.
32. Haghayegh S, Kang H-A, Khoshnevis S, et al. A comprehensive guideline for Bland–Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings. *Physiol Measur.* 2020;41:055012.
33. Giavarina D. Understanding bland altman analysis. *Biochem Medica.* 2015;25:141–151.
34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropract Medicine.* 2016;15:155–163.
35. Wittig L, Betz C, Eggert D. Optical coherence tomography for tissue classification of the larynx in an outpatient setting-a translational challenge on the verge of a resolution? *Transl Biophoton.* 2021;3:e202000013.
36. Benboujja F, Greenberg M, Nourmahnad A, et al. Evaluation of the human vocal fold lamina propria development using optical coherence tomography. *The Laryngoscope.* 2021;131:E2558–E2565.

37. Kobler JB, Chang EW, Zeitels SM, Yun SH. Dynamic imaging of vocal fold oscillation with four-dimensional optical coherence tomography. *The Laryngoscope.* 2010;120:1354–1362.
38. Zhang Z. Vocal fold vertical thickness in human voice production and control: a review. *J Voice.* 2025;39:1183–1191. https://doi.org/10.1016/j.jvoice.2023.02.021.Epub 2023 Mar 22. PMID: 36964073; PMCID: PMC10514229.
39. Zhang Z. Contribution of laryngeal size to differences between male and female voice production. *J Acoust Soc Am.* 2021;150:4511–4521.
40. Wu L, Zhang Z. Voice production in a MRI-based subject-specific vocal fold model with parametrically controlled medial surface shape. *J Acoust Soc Am.* 2019;146:4190–4198.
41. Zhang Z. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *J Acoust Soc Am.* 2016;139:1493–1507.
42. Lohscheller J, Švec JG, Döllinger M. Vocal fold vibration amplitude, open quotient, speed quotient and their variability along glottal length: Kymographic data from normal subjects. *Logoped Phoniatr Vocol.* 2013;38:182–192.
43. Titze IR. Physiologic and acoustic differences between male and female voices. *J Acoust Soc Am.* 1989;85:1699–1707.
44. Bonilha HS, Deliyski DD, Gerlach TT. Phase asymmetries in normophonic speakers: visual judgments and objective findings. *Am J Speech-Lang Pathol.* 2008;17:367–376.