Article

Direct Measurement and Modeling of Intraglottal, Subglottal, and Vocal Fold Collision Pressures during Phonation in an Individual with a Hemilaryngectomy

Daryush D. Mehta [1,2,3,4,*], James B. Kobler [1,2,3], Steven M. Zeitels [1,2,3], Matías Zañartu [5], Emiro J. Ibarra [5], Gabriel A. Alzamendi [6], Rodrigo Manriquez [5], Byron D. Erath [7], Sean D. Peterson [8], Robert H. Petrillo [1] and Robert E. Hillman [1,2,3,4]

1 Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA 02114, USA; james.kobler@mgh.harvard.edu (J.B.K.); zeitels.steven@mgh.harvard.edu (S.M.Z.); rhpetrillo@mgh.harvard.edu (R.H.P.); hillman.robert@mgh.harvard.edu (R.E.H.)
2 Department of Surgery, Massachusetts General Hospital–Harvard Medical School, Boston, MA 02114, USA
3 Speech and Hearing Bioscience and Technology, Division of Medical Sciences, Harvard Medical School, Boston, MA 02115, USA
4 MGH Institute of Health Professions, Boston, MA 02129, USA
5 Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile; matias.zanartu@usm.cl (M.Z.); emiro.ibarra@sansano.usm.cl (E.J.I.); rodrigo.manriquezp@usm.cl (R.M.)
6 Institute for Research and Development on Bioengineering and Bioinformatics, National University of Entre Rios–CONICET, Entre Ríos 3100, Argentina; galzamendi@ingenieria.uner.edu.ar
7 Department of Mechanical & Aeronautical Engineering, Clarkson University, Potsdam, NY 13699, USA; berath@clarkson.edu
8 Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; peterson@uwaterloo.ca
* Correspondence: mehta.daryush@mgh.harvard.edu

Featured Application: The overall goal of this work is to better understand how vocal fold collision contributes to the development and clinical management of vocal pathologies, such as vocal fold nodules and polyps, and to ultimately develop measures that will improve the prevention, diagnosis, and treatment of phonotraumatic voice disorders.

Abstract: The purpose of this paper is to report on the first in vivo application of a recently developed transoral, dual-sensor pressure probe that directly measures intraglottal, subglottal, and vocal fold collision pressures during phonation. Synchronous measurement of intraglottal and subglottal pressures was accomplished using two miniature pressure sensors mounted on the end of the probe and inserted transorally in a 78-year-old male who had previously undergone surgical removal of his right vocal fold for treatment of laryngeal cancer. The endoscopist used one hand to position the custom probe against the surgically medialized scar band that replaced the right vocal fold and used the other hand to position a transoral endoscope to record laryngeal high-speed videoendoscopy of the vibrating left vocal fold contacting the pressure probe. Visualization of the larynx during sustained phonation allowed the endoscopist to place the dual-sensor pressure probe such that the proximal sensor was positioned intraglottally and the distal sensor subglottally. The proximal pressure sensor was verified to be in the strike zone of vocal fold collision during phonation when the intraglottal pressure signal exhibited three characteristics: an impulsive peak at the start of the closed phase, a rounded peak during the open phase, and a minimum value around zero immediately preceding the impulsive peak of the subsequent phonatory cycle. Numerical voice production modeling was applied to validate model-based predictions of vocal fold collision pressure using kinematic vocal fold measures. The results successfully demonstrated feasibility of in vivo measurement of vocal fold collision pressure in an individual with a hemilaryngectomy, motivating ongoing data collection that is designed to aid in the development of vocal dose measures that incorporate vocal fold impact collision and stresses.

1. Introduction

Voice disorders have been estimated to affect approximately 30% of the adult population in the United States at some point in their lives, with up to 7.6% of individuals affected at any given point in time [1,2]. The most common voice disorders are chronic or recurring conditions believed to result from excessive and/or poorly regulated activity of the perilaryngeal muscles, referred to as vocal hyperfunction. Vocal hyperfunction can be associated with either phonotrauma-induced lesions of the vocal folds (e.g., nodules and polyps) or with dysphonia occurring in the absence of vocal fold trauma or structural or neurological abnormalities (e.g., primary muscle tension dysphonia) [3]. The etiology of phonotraumatic lesions has classically centered on the role of vocal fold impact stress (in the direction of tissue motion at contact) and shear stress (along the tissue surface) during phonation [4]. Insight into the role of collision pressures in the generation of vocal fold lesions is important for any quantification or modeling of phonotrauma since phonotraumatic lesions are widely considered to form due to repetitive stress on the mid-membranous portion of the vocal folds at the location of maximal tissue contact [5–7].

The direct measurement of vocal fold collision pressure is challenging to carry out in vivo, and only two published studies have attempted to gather data from sensors placed intraglottally during phonation [8,9]. Verdolini et al. [9,10] developed a piezoresistive pressure sensor with a flat frequency response up to 50 kHz and a linear sensing range of 0–14 kPa (0–140 cm H_2O). The circular pressure-sensing element was 1.8 mm in diameter (0.4 mm thickness) and inserted transorally via a curved cannula. Limited in vivo data were successfully obtained from vocally healthy individuals given challenges related to sensor positioning between two vibrating vocal folds without disrupting phonation and uncertainty as to whether the pressure signal was related to vocal fold collision or acoustic (non-impact) pressures. Vocal fold collision pressures across study participants were reported to be in the range of 4–32 cm H_2O. However, the criteria by which collision pressure signals in that study were deemed to be of adequate quality were unclear. Only one exemplary waveform exhibited peaks in the pressure signal at the expected vocal fold oscillation rate, but noise in the signal (and a flat-line microphone signal) prevented the ability to confirm whether the peaks were due to impact stress sensing or subglottal (or supraglottal) acoustic pressure sensing.

Gunter et al. [8] employed a piezoelectric, force-sensitive plate (10 mm $\times$ 15 mm for a sensing area of 150 mm^2; 0.29 mm thickness) at the end of a transorally positioned rigid cannula that was inserted between the vocal folds. The sensing element exhibited a flat frequency response up to 25 kHz and a linear sensing range for measured forces from a 2.5 mN noise floor up to 200 mN (16–1333 Pa, or 0.16–13.6 cm H_2O). In that study, peaks in the electroglottography signal and force sensor signal provided an indication of whether the force peaks measured were due to air pressure (force sensor peak lagging behind the electroglottography peak) or vocal fold collision forces (force sensor peak aligned with the electroglottography peak). Intraglottal placement of the force sensor during phonation was considered satisfactory when endoscopic imaging confirmed the desired sensor positioning against one vocal fold, anterior–posterior positioning in the mid-membranous glottis, and inferior–superior positioning of the force sensor in the coronal plane of vocal fold contact (the strike zone). The exemplary waveforms in that paper displayed a force sensor output that resembled a triangular waveform that has not been observed in computational or bench-top models of phonation. The authors acknowledged that the force sensor signal could also reflect acoustic or aerodynamic pressure variations due to the size of the sensor extending above and below the plane of vocal fold contact.

Both of these in vivo studies demonstrated the critical role of endoscopic image guidance to aid in the verification of intraglottal sensor placement. In practice, these

types of experiments require skillful ambidexterity to position two probes at the same time (one hand holding the intraglottal sensor probe and the other holding the imaging endoscope) and to accurately position the intraglottal sensor with minimal tactile feedback from the laryngeal tissue while viewing a two-dimensional video image; laryngologists develop these skills to perform office-based transoral vocal fold injections to treat vocal fold paralysis and other pathologies. Equally challenging is having to take into account that the coronal plane of vocal fold collision is higher (more superior) during phonation than when the vocal folds are at rest during breathing. Relying solely on visual feedback potentially misses capturing the true intraglottal pressure signal during phonation. Multichannel techniques using data from multiple pressure sensors can aid in the real-time monitoring of intraglottal positioning, e.g., multi-channel electroglottography applies a similar concept to monitor and verify correct electrode positioning at the glottal level in real time [11]. Essentially, if the signals captured by two adjacent pressure sensors are in phase and correlated, then they are both positioned in the same location—supraglottally, subglottally, or intraglottally. If the two sensor signals have waveshapes distinct from each other, their difference signal, polarity, and degree of correlation help confirm that one of the sensors is positioned as desired between the glottis and in the strike zone of collision during phonation [12,13].

Expectations with respect to the waveform shape of the intraglottal pressure signal come from numerical models of phonation [14–20], self-oscillating physical models of synthetic vocal fold-like material [12,14,21], aerodynamically driven excised larynges or hemilarynx models [13,22–28], and in vivo animal models [29]. For example, in a hemilarynx model, the superior–inferior position of a pressure sensor was systematically varied as the model was driven into self-sustained oscillation with an external airflow source [13]. The position of the sensor was confirmed to be in the strike zone of vocal fold collision using a dual-imaging technique that provided en face imaging of the medial vocal fold surface and top-down imaging of the vibrating vocal fold. Taken together with the other modeling results, this body of literature provides strong evidence that the in vivo intraglottal pressure signal during phonation should have three primary components: (1) an impulsive peak in the direction of increasing pressure at the start of the phonatory closed phase (collision/impact component), which is followed in time by (2) a more rounded peak during the phonatory open phase (acoustic/aerodynamic pressure component), and (3) a minimum value around zero immediately preceding the impulsive peak of the subsequent phonatory cycle. To date, these expected characteristics of the intraglottal pressure waveform have not been directly observed in human speakers.

Direct measurement of vocal fold collision pressures during bilateral vocal fold phonation has proven to be challenging in vivo [8,9]. Positioning a sensor against one vocal fold in a typical larynx can result in disrupted vocal fold vibration and irregular glottal closure characteristics. However, these issues can be mitigated in a group of individuals who have undergone surgical treatment for laryngeal cancer but still exhibit functional voice outcomes due to the cancer being largely limited to one vocal fold [30–32]. This treatment involves performing an endoscopic hemilaryngectomy to remove the primarily affected vocal fold, while preserving as much as possible of the healthy tissue and pliability of the superficial lamina propria on the contralateral side (and, thus, preserve the potential for vocal function). The side with the vocal fold removed is then surgically reconstructed to form a scar band that is medialized using implantable materials (e.g., adipose tissue) so that the vibrating contralateral vocal fold can achieve glottal closure during phonation. The reconstructed larynx is analogous to the excised hemilarynx setup that has been previously employed to study phonatory physiology, e.g., [13,33–35]. Many of these patients have perceptually typical conversational voices that are much improved compared with their pre-surgical condition, when they often exhibit insufficient glottal closure that produces excessively inefficient voice production and a breathy voice quality [36].

The primary aim of this paper is to report on the first in vivo application of a recently developed transoral, dual-sensor intraglottal/subglottal pressure (ISP) probe [12,13] in

such an individual with a hemilaryngectomy. In these individuals, the tip of the ISP probe is designed to rest against the surgically medialized scar band to yield stable and reliable collision pressure measurements of the medial surface of the contralateral vocal fold (essentially acting as an excised hemilarynx configuration). The probe was designed to simultaneously measure the intraglottal pressure signal from a proximal sensor and subglottal pressure from a second, distal sensor at the tip of the probe to gain empirical insight into in vivo vocal fold collision and aerodynamic relationships during phonation. When positioned appropriately during phonation, the intraglottal pressure signal is expected to consist of the two expected components of an impulsive peak due to vocal fold collision and a more rounded peak due to intraglottal aerodynamic pressures during the open phase of the phonatory cycle.

A secondary aim of this paper is to illustrate how numerical models of phonation can be validated using the in vivo measurement of vocal fold collision pressures as reference values. Numerical models allow clinicians and researchers to derive hard-to-measure parameters related to voice production (such as vocal fold collision pressures) from relatively easier-to-measure parameters (such as glottal area waveforms) since invasive direct measurements are generally not feasible in practice. Two numerical models are evaluated that enable the derivation of vocal fold collision pressures from high-speed videoendoscopic imaging data.

2. Materials and Methods

2.1. Participant Characteristics

A 78-year-old male was recruited for this study who had previously (12 years prior) undergone a unilateral, right vocal fold hemilaryngectomy for the treatment of laryngeal cancer (T2bN0M0 squamous cell carcinoma). An adipose tissue implant, taken from the periumbilical region, was injected to medialize the right vocal fold and facilitate glottal closure as the left vocal fold oscillated during phonation. The individual follows up regularly with his treating laryngologist for cancer surveillance and to assess vocal function, which includes laryngeal videoendoscopic examinations with stroboscopy. Figure 1 displays endoscopic images of the individual's larynx in an abducted and adducted state. Video S1 (Supplementary Materials) displays a segment from the videostroboscopic examination as the individual produces voice at a comfortable and higher-than-comfortable pitch level. According to clinical notes, the individual exhibited normal pliability and mucosal wave of the left vocal fold and diminished pliability and mucosal wave of the right vocal fold, consistent with expectations following the hemilaryngectomy and reconstruction. Complete glottal closure was observed during phonation with entrained vocal fold oscillation, mild phase asymmetry, and significantly reduced amplitude of the right vocal fold. There was no evidence of recurrent laryngeal disease.

(a)   (b)
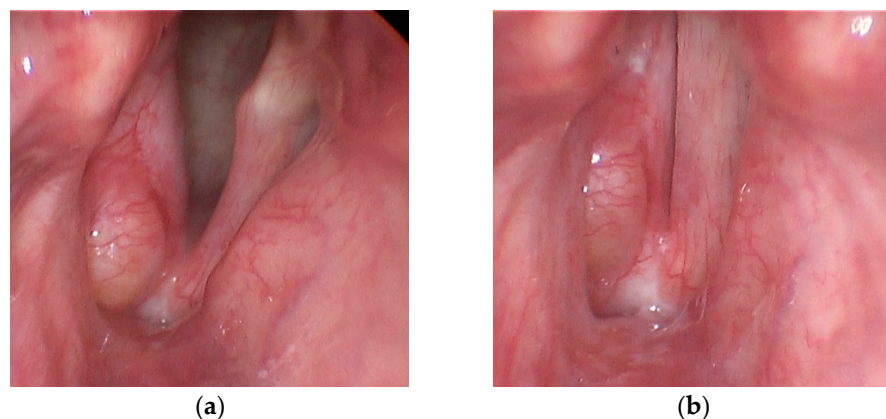
Figure 1. Endoscopic images of the larynx from the videostroboscopy examination of the participant who had previously undergone a unilateral (right vocal fold) hemilaryngectomy to treat laryngeal cancer. Shown are snapshots of the vocal folds in states of (**a**) abduction and (**b**) adduction.

2.2. Data Collection

Following the clinical videostroboscopic assessment, the participant underwent a laryngeal high-speed videoendoscopic assessment with simultaneously recorded vocal function sensors. The endoscopist (a laryngologist who regularly performs outpatient transoral vocal fold injections) positioned the ISP probe with the left hand against the scar band on the right side of the larynx and synchronously recorded laryngeal high-speed videoendoscopy via a 70-degree transoral rigid endoscope (10 mm outer diameter; JEDMED, St. Louis, MO, USA) held in the right hand. The high-speed camera was attached to the eyepiece of the endoscope (FASTCAM Model MC2, Photron, Tokyo, Japan) using a 45 mm focal length lens adapter (PENTAX Medical, Montvale, NJ, USA) and high-speed videoendoscopy data were recorded with a Color High-Speed Video System (Model 9710, PENTAX Medical, Montvale, NJ, USA). The high-speed video data were recorded at 2000 frames per second with maximum frame integration time (~0.5 ms) and a spatial resolution of 512 pixels $\times$ 512 pixels. The imaging light source consisted of a 300-watt short-arc xenon lamp.

Five channels of sensor data recorded by an acoustic microphone, electroglottograph, neck-surface accelerometer, and the two pressure sensors of the ISP probe were time-synchronized with the high-speed video data. The acoustic signal was recorded 15 cm from the lips using a head-mounted, omnidirectional microphone (model ME102, Sennheiser Electronic GmbH, Wedemark-Wennebostel, Germany). Electroglottography (Model EG2-PC; Glottal Enterprises, Syracuse, NY, USA) provided a signal associated with vocal fold tissue contact that is high-pass filtered to reduce noise and low-frequency conductance changes. Neck-surface vibration was measured using an accelerometer (BU-27135; Knowles, Itasca, IL, USA) that can be analyzed to yield glottal airflow estimates [37,38] and has shown potential for long-term monitoring of vocal function, e.g., [39–42]. The accelerometer was enclosed in a silicone epoxy and mounted onto the participant's neck halfway between the sternal notch and thyroid prominence using Double Stick Discs (Model 2181, 3M, Maplewood, MN, USA).

Figure 2 displays a photograph of the dual-sensor ISP probe and its dimensions [13]. The width of the probe tip was 4.4 mm, and the probe tip length (before curvature began) was 28 mm. The thickness of the probe at the sensor locations was maximally 1.9 mm, which is larger than in previous devices (0.4 mm [9,10] and 0.29 mm [8]) due to the dimensions of the pressure-sensing elements. The effect of the thickness of the probe tip was mitigated by the placement of the probe against the surgically reconstructed vocal fold (scar band) of the study participant as the contralateral vocal fold came into contact with the pressure probe. The two probe catheters were connected, via extender adapter cables, to a two-channel signal conditioning unit (PCU-2000, Millar, Inc., Houston, TX, USA) that provides electrical isolation, analog knobs for zero control, and a flat frequency response of 0–1000 Hz. Measurement of intraglottal and subglottal pressures was accomplished using the two miniature, pressure-sensitive sensors of the ISP probe, respectively (Mikro-Cath Pressure Catheter, Millar, Inc., Houston, TX, USA). The pressure-sensing elements consisted of diffused, piezoresistive semiconductors with a dynamic range of -6.7 kPa to 40 kPa (-70 cm H_2O to 400 cm H_2O) and a flat frequency response up to 10 kHz. Each pressure transducer consists of an ovoid capsule that is 4.8 mm long and 1.17 mm in diameter, connected to a 120 cm long flexible cable. The sensing element measured 1.0 mm $\times$ 1.0 mm (sensing area of 1.0 mm^2) and was recessed 0.37 mm into the cylindrical surface of the transducer [12]. To reduce the uncertainty in collision pressure measurements by the ISP probe, the sensors were embedded in medical-grade room-temperature-vulcanizing (RTV) silicone and covered with a thin 0.125 mm silicone sheet to create a flat contact surface [12,13].

The five data channels were each digitized at 8000 samples per second and 16-bit quantization by a digital acquisition board (6259 M series, National Instruments, Austin, TX, USA). Gain control and anti-aliasing filtering was set by preconditioning electronics (CyberAmp model 380, Danaher, Corp., Washington, DC, USA). Time synchronization of

the high-speed video data and data channels was accomplished by a common clock source from the National Instruments board that supplied the 2000 Hz (video) and 8000 Hz (data) sampling signals.
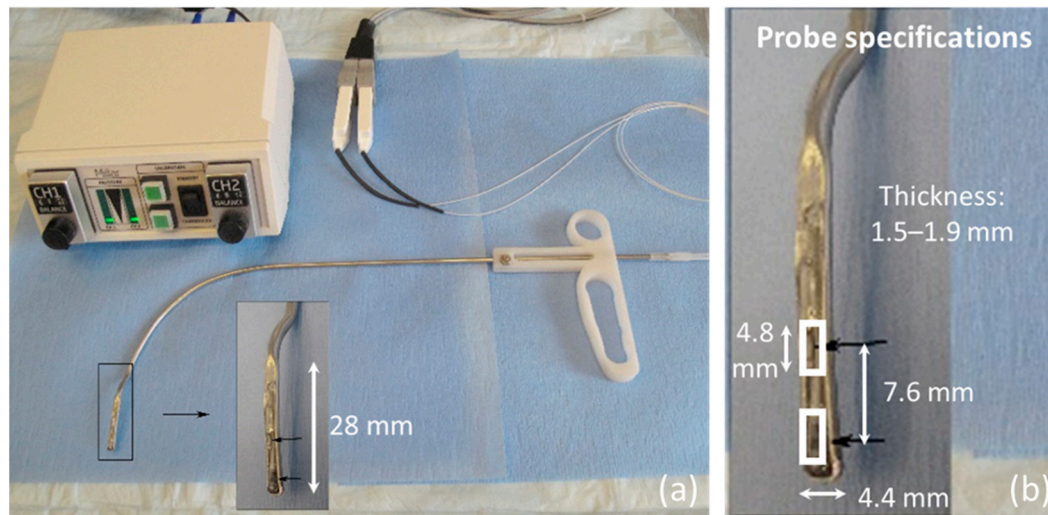

Figure 2. In vivo intraglottal/subglottal pressure (ISP) probe with two pressure sensors at the probe tip to simultaneously measure intraglottal and subglottal pressure during phonation, (**a**) ISP probe with a Ford injector-like handle and two-channel signal conditioning electronics, (**b**) zoomed-in view of the ISP probe tip showing dimensions of the two in-line pressure sensors. From [13].

Two of the alternating current (AC) signals were calibrated using linear scaling factors to convert from voltage levels to units of pascals for the microphone signal [43] and units of cm/s^2 for the neck-surface accelerometer signal. The AC electroglottography signal was left uncalibrated as a relative measure vocal fold contact. Calibration of the ISP probe's pressure sensors accounted for the sensor silicone enclosures by submerging the probe in a graduated cylinder filled with water and noting the voltage level at given submergence depths for each sensor (i.e., hydrostatic pressure) and computing a best-fit line to the data (coefficient of determination for the line was 1.0). The linear, multiplicative scale factors that mapped pressure sensor voltage levels (V) to units of pressure (cm H_2O) were stable and repeatable at 1.596 (cm H_2O)/V for the intraglottal pressure sensor and 1.722 (cm H_2O)/V for the subglottal pressure sensor. The DC level of the intraglottal and pressure signals exhibited a low-frequency component (DC drift) also observed in prior in vivo studies that was associated with thermal fluctuations [8,9]. Zero subglottal pressure was defined during the silence period prior to the onset of phonation (3.4 cm H_2O correction factor was subtracted from the measured signal), and the intraglottal zero pressure level was defined at the most-negative peak pressure value [13] when the intraglottal sensor was positioned in the plane of vocal fold collision during phonation (1.3 cm H_2O correction factor was added to the measured signal).

Endoscopic visualization of the larynx during sustained phonation at comfortable pitch and loudness levels allowed the endoscopist to place the dual-sensor pressure probe such that the proximal sensor was positioned intraglottally and the distal sensor subglottally. The clinician held a left-handed ISP probe with pressure-sensing elements facing the functioning left vocal fold of the participant. As recommended in prior work [13], it was necessary to ask the subject to produce sustained phonation while the ISP probe was slowly swept in the superior–inferior dimension. Then, during the data analysis phase, features related to vocal fold collision were then identified to determine when adequate vocal fold contact occurred to capture peak collision pressures. Figure 3 displays high-speed videoendoscopic images from one phonatory cycle during which the proximal pressure sensor of the ISP probe was deemed to be positioned in the strike zone of vocal fold collision. Video

S2 (Supplementary Materials) shows the corresponding high-speed videoendoscopy data for this phonatory segment.
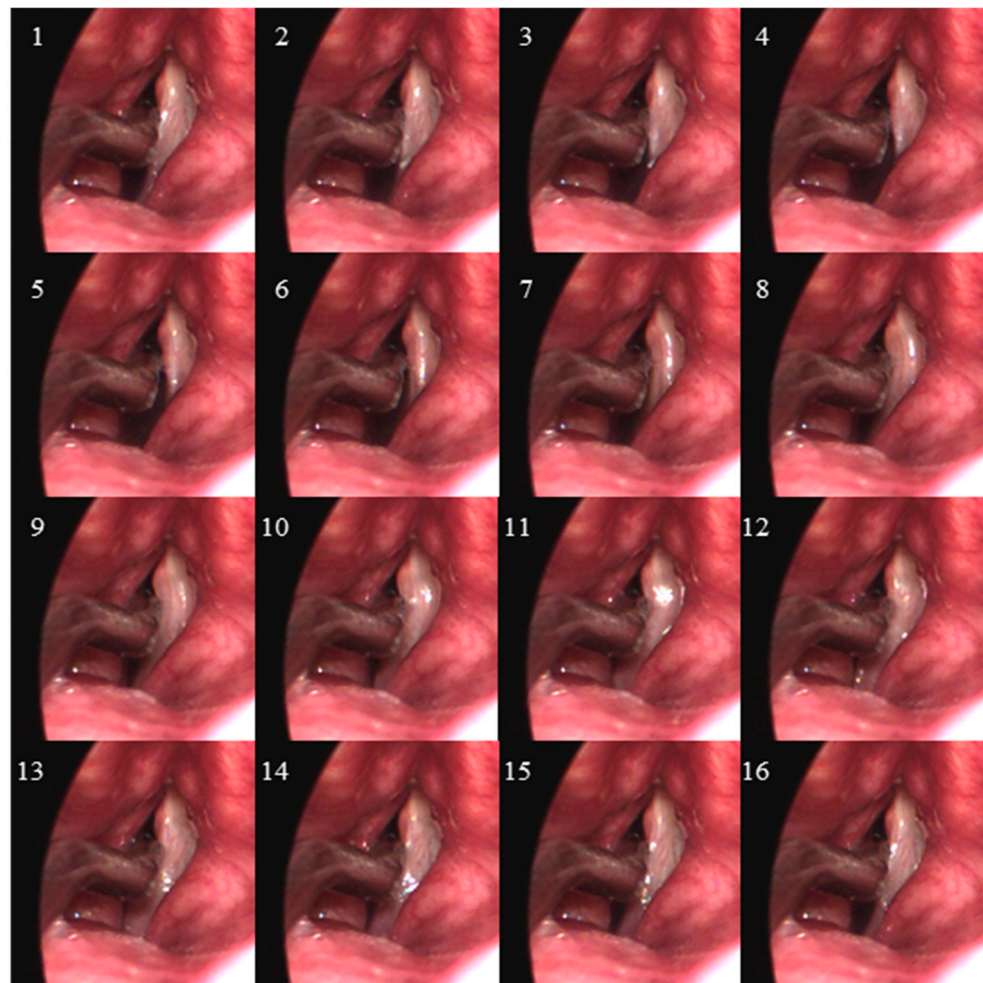
Figure 3. One phonatory cycle is displayed from the high-speed videoendoscopy data (2000 frames per second) during which the intraglottal pressure sensor of the ISP probe was deemed to be in the strike zone of the left vocal fold. Sixteen frames (frame indices indicated) are shown for the phonatory cycle that was 8 ms in duration, translating to a fundamental frequency of tissue oscillation of 125 Hz.

2.3. Data Analysis

The proximal pressure sensor of the ISP probe was considered to be in the strike zone of vocal fold collision when the intraglottal pressure signal exhibited the three expected characteristics observed in the literature (e.g., [13]): an impulsive peak at the start of the closed phase, rounded peak during the open phase, and minimum value around zero immediately preceding the impulsive peak of the subsequent phonatory cycle. For phonatory segments exhibiting vocal fold collision, the peak collision pressure was defined as the maximum pressure value for each cycle. The mean subglottal pressure was computed from the simultaneously recorded signal of the distal pressure sensor of the ISP probe. The fundamental frequency of the intraglottal pressure waveform was computed using an autocorrelation-based method designed to process acoustic voice signals [44]. Vocal sound pressure level was computed from the root mean square of the calibrated acoustic microphone signal in dB SPL at 15 cm from the lips.

Two numerical modeling approaches were applied to validate their estimates of vocal fold collision pressure given the glottal area waveform and/or vocal fold kinematics derived from the high-speed videoendoscopy data. The first numerical approach employed a Hertzian impact model to estimate vocal fold collision pressures using vocal fold edge

and glottal area contour analysis, referred to as contact pressure analysis (see [45] for algorithmic details). This approach took vocal fold kinematic information extracted from laryngeal high-speed videoendoscopic data as input only. A Kalman filter scheme linked the vocal fold edge motion with a lumped-mass representation of vocal fold tissue contact mechanics. The lumped-mass model was used to estimate the non-physical overlap in the Hertz contact model associated with the deformation of two colliding cylinders. For both consistency and simplicity, the vibrating tissue of each vocal fold edge was represented by a parabola, with anterior and posterior anchor points that were selected for each vocal fold. Given that the tip of the ISP probe rested on the right vocal fold of the participant in this study (blocking the visualization of the vocal fold; see Figure 3), manual adjustment of the anchor points was performed such that the vibrating tissue of the left vocal fold collided at the probe location. Even though this contact pressure analysis method was originally designed to produce normalized contact pressure values, results in physical units were obtained by applying generic material properties for the vocal fold tissue, an effective Young's modulus of 24.75 kPa, and a Poisson ratio of 0.5 [45]. Note that this approach only provides point estimates of the peak vocal fold collision pressure. Alternative models are necessary to yield confidence intervals on these point estimates, as well as estimates of other important vocal function measures.

The second mathematical model accomplished the goal of deriving vocal fold collision pressure and other hard-to-measure physiological measures, such as subglottal pressure, glottal airflow, and intrinsic muscle activation levels. This model utilized a Bayesian estimation technique to analyze the spatially calibrated glottal area waveform to derive time-varying signals and confidence intervals for vocal fold collision pressure and these other physiological measures (see [46] for algorithmic details). This approach was based on a subject-specific, body-cover model of the vocal folds [47] and an extended Kalman filter with the glottal area waveform as the only observation signal, i.e., Case I in Alzamendi et al. [46]. The glottal area was obtained using the Glottal Image Explorer software [48], which also allowed for estimating the portion of the glottal area that was visually blocked by the ISP probe by adjusting the software parameters. To mimic the hemilaryngectomy condition of the left vocal fold coming into contact with the ISP probe, only the left vocal fold was analyzed, and a symmetric condition was assumed. Spatial calibration of the high-speed video was accomplished using the known dimensions of the probe tip as an imaging ruler. Model outputs were subglottal pressure, posterior glottal gap area, intrinsic muscle activation levels of the cricothyroid and thyroarytenoid, and vocal fold collision pressure. Each of the outputs was computed with 95% confidence intervals. Initial model conditions were set to yield a large initial uncertainty at the beginning of the signal to assess the speed of model convergence. For both models, the error was computed between the model-estimated peak vocal fold collision pressure and measured collision pressure from the ISP probe data.

3. Results

The intraglottal pressure sensor was deemed to be positioned as desired in the phonatory strike zone by comparing the intraglottal pressure sensor waveform with that of the subglottal pressure sensor. If the two sensor signals were positively correlated and in phase with each other, then both sensors were determined to both be positioned either subglottally or supraglottally. If the two sensor signals had significant waveform differences between them (with cycles occurring at the same fundamental frequency), then the two sensors were considered to be (desirably) in different locations within and around the glottis. The waveform of the (proximal) intraglottal pressure sensor was further investigated to determine if the expected waveshape characteristics, per the literature, were observed.

3.1. Direct Measurement of Aerodynamic and Acoustic Signals

Figure 4 displays the synchronized data signals during a 100 ms phonatory segment during which vocal fold collision was sensed by the intraglottal pressure signal. The vocal

sound pressure level was 81.4 dB SPL at 15 cm, fundamental frequency was 126.1 Hz, mean subglottal pressure was 9.0 cm H_2O, and the mean peak collision pressure was 9.0 cm H_2O. The ratio between the peak collision pressure and subglottal pressure was thus 1.0, which is within the range of 0.5 to 3.8 that has been observed in excised hemilarynx experiments [13,22]. We believe this is the first time that the expected intraglottal pressure waveform has been captured in vivo in a human speaker with the two expected waveform components due to impact pressure at the start of the closed phase (impulsive positive-going peak) and aerodynamic pressure during the open phase (more rounded positive-going peak). The conductance signal of the electroglottograph corroborates the timing of the open and closed phases in relation to the intraglottal/collision pressure signal. The polarity of the accelerometer was such that motion away from the body (perpendicular to the neck surface) was recorded as positive-going amplitude values. The microphone signal captured the negative-going impulse per cycle that is associated with the excitation of the voice source by the maximum glottal flow declination rate [49]. Video S3 (Supplementary Materials) shows an excerpt of this phonatory segment in a graphical user interface that displays the synchronized high-speed videoendoscopy with the multimodal sensor signals from the acoustic microphone, electroglottograph, neck-surface accelerometer, and the two ISP probe signals of intraglottal and subglottal pressure. The sensor signals are provided in WAV format in the Supplementary Materials section as Signals S5–S9.
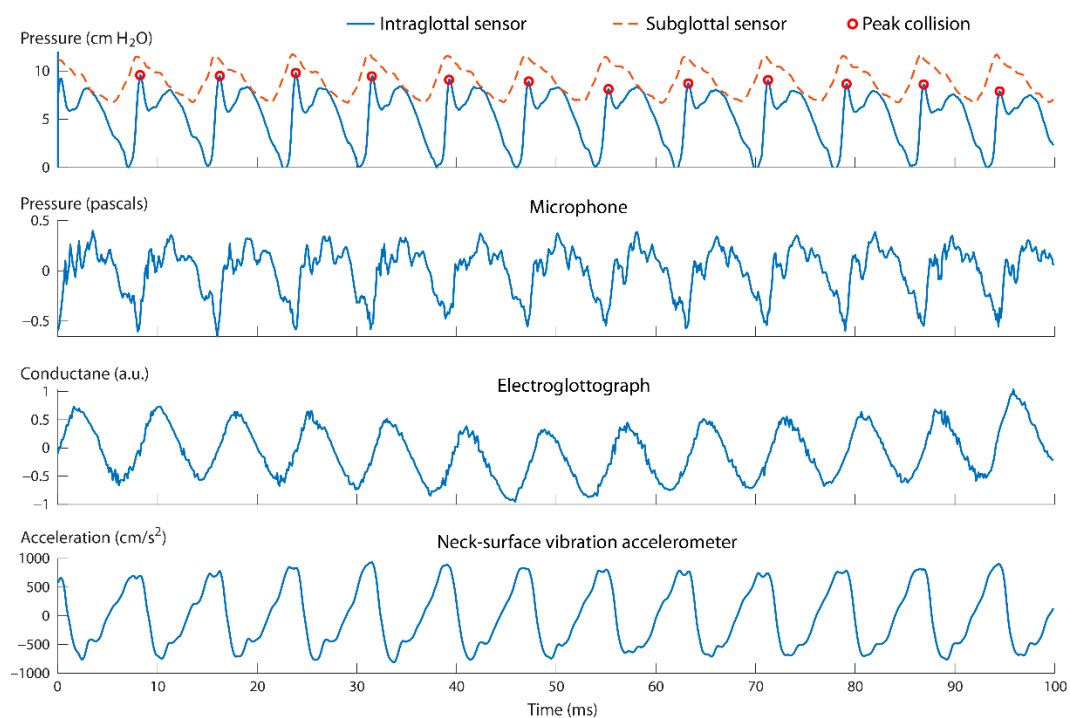

Figure 4. Calibrated signals recorded during a phonatory segment when the intraglottal pressure sensor was determined to be in the strike zone. Shown as circles are the peak collision pressures in the intraglottal sensor signal during each phonatory cycle.

3.2. Validation of the Two Model-Based Estimates of Vocal Fold Collision Pressure

Figure 5 shows the results of the contact pressure analysis method that incorporated a Hertzian model of collision [45]. Plotted are the input measured variables to the model from the high-speed videoendoscopy imaging data (calibrated glottal area and vocal fold displacement) and output estimated parameters from the model (vocal fold deformation at contact and vocal fold collision pressure). See Video S4 (Supplementary Materials) showing a movie of Figure 5 that plots over time the input and output variables overlaid on the high-speed videoendoscopy frames. Direct comparison between the estimates of collision pressure from the model and the measured intraglottal pressure sensor signal

is illustrated in Figure 5 as well. The impact stress component matches the first peak of the measured intraglottal signal well. The root mean square error between the estimated collision pressure peaks and the measured peak collision pressures across all cycles in this 100 ms phonatory segment was 1.04 cm H_2O, corresponding to a mean absolute error of 0.81 cm H_2O and mean absolute percentage error of 9.1%. The resulting contact pressure component is in good agreement with prior numerical studies [18].
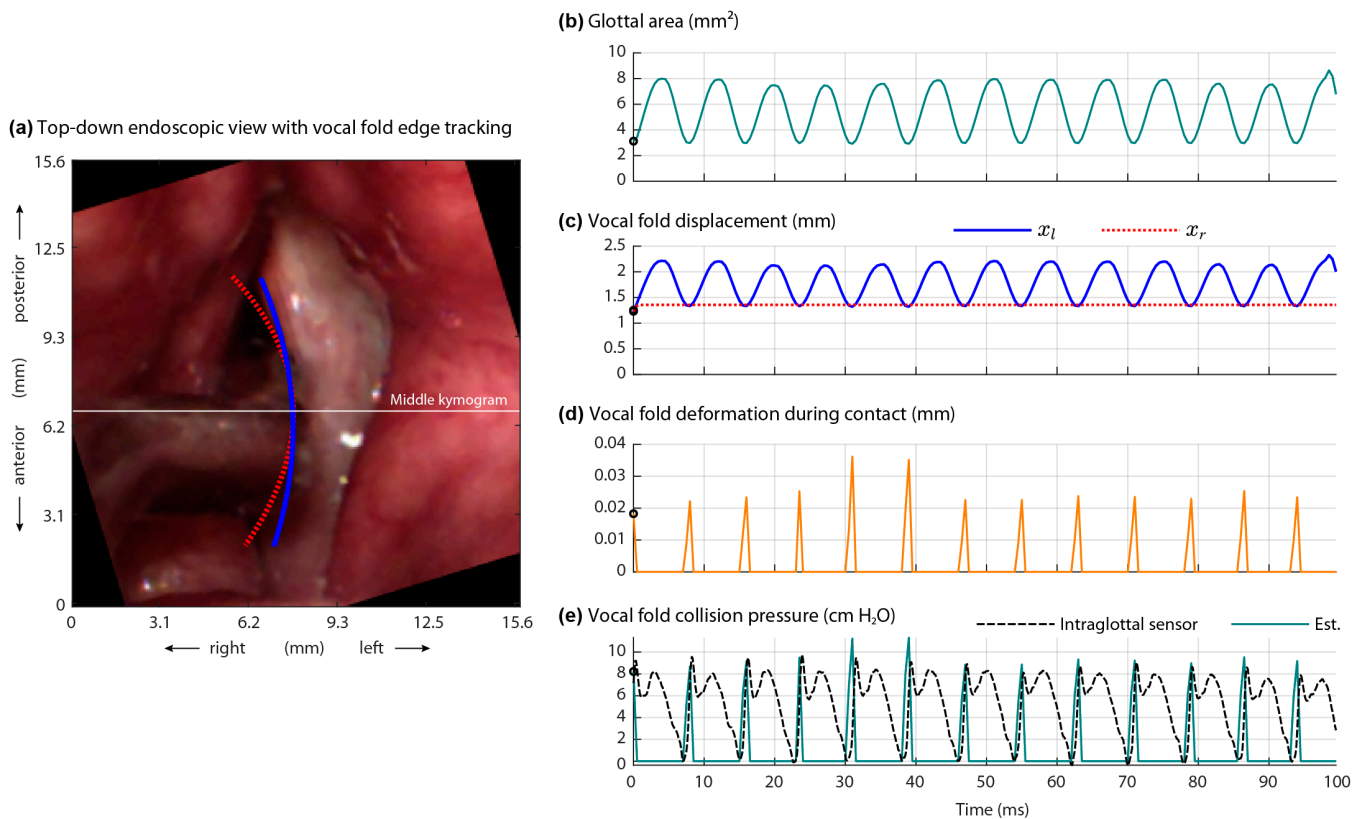
Figure 5. Results of the Hertzian contact pressure analysis method of Díaz-Cádiz et al. [45]. Shown are the (**a**) first high-speed videoendoscopic frame with vocal fold edge tracking, (**b**) measured glottal area waveform, (**c**) vocal fold displacement at the middle digital kymogram of the left (x_l) and right (x_l) folds using a Kalman filter, (**d**) vocal fold deformation during contact (model output), and (**e**) vocal fold collision pressure as estimated by the model (Est.) and directly measured by the ISP probe's intraglottal sensor.

Figure 6 shows the results of the second model that applied Bayesian estimation to a subject-specific, lumped-mass vocal fold model [46]. Plotted on the left panels of Figure 6 are the input measured variables to the model from the high-speed videoendoscopy imaging data: calibrated glottal area, vocal fold displacement, and vocal fold velocity of the medial edge. On the right panels of Figure 6 are the output estimated parameters from the model: normalized muscle activation levels of the cricothyroid and thyroarytenoid muscles (a_{CT} and a_{TA}, respectively), subglottal pressure, and the vocal fold collision pressure signal. In addition to the parameter estimates, Bayesian estimation also provides the confidence intervals associated with each output parameter waveform. The large initial model uncertainty within the first 10 ms was expected due to the initial conditions applied; following that transient period, the confidence interval of the model outputs rapidly decreased and reached steady-state values, illustrating good model convergence with tight confidence intervals. As found in direct measurement, an approximate 1:1 relation between the model-based peak collision pressure and subglottal pressure values was observed using the Bayesian estimation method. The root mean square error between the estimated vocal fold collision pressure peaks and the measured peak collision pressures was 1.82 cm H_2O,

corresponding to a mean absolute error of 1.58 cm H_2O and mean absolute percentage error of 17.9%.
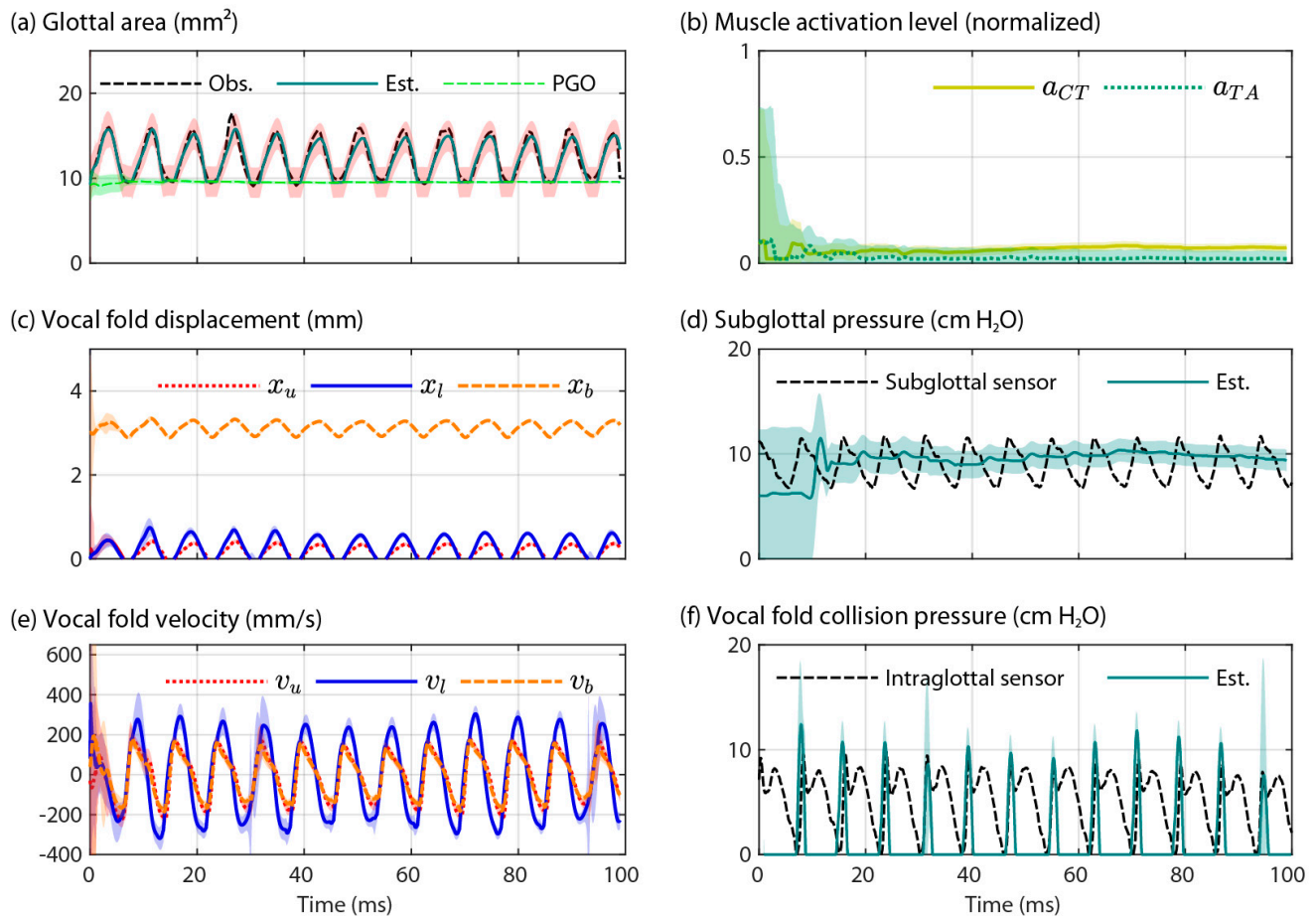
Figure 6. Bayesian estimation of physiological parameters of interest using the subject-specific modeling approach of Alzamendi et al. [46]. Shown are the (**a**) observed (Obs.) and estimated (Est.) glottal area waveforms with estimated posterior glottal opening (PGO), (**b**) estimated muscle activation levels of the cricothyroid (a_{CT}) and thyroarytenoid muscles (a_{TA}), (**c**) measured vocal fold displacement for the upper (x_u), lower (x_l), and body (x_b) masses of the left (vibrating) vocal fold, (**d**) subglottal pressure as estimated by the model (Est.) and directly measured by the ISP probe's subglottal pressure sensor, (**e**) measured vocal fold velocity for upper (x_u), lower (x_l), and body (x_b) masses of the left (vibrating) vocal fold, and (**f**) vocal fold collision pressure as estimated by the model (Est.) and directly measured by the ISP probe's intraglottal pressure sensor. Shaded areas correspond to 95% confidence intervals around the model-estimated parameters.

4. Discussion

This study was the next step in ongoing work to demonstrate the feasibility of obtaining direct measurements of the vocal fold collision pressure during phonation in an individual that was specially selected to reduce the challenges of performing the procedure in a typical individual with bilaterally vibrating vocal folds. The synchronized collection of estimates of subglottal air pressure below the vocal folds was necessary to effectively interpret the impact of aerodynamic versus tissue collision forces on the pressure measurements being obtained at the vocal fold level. The simultaneous recording of high-speed videoendoscopy and non-invasive recordings of neck-surface accelerometry, acoustics, and electroglottography allow for the cross-correlation of these measures with the intraglottal/subglottal pressure signals of the ISP probe. Such data are much more valid and valuable if they can be reliably acquired in vivo as opposed to the excised situation where there is no neural innervation of associated flaccid muscles and no perfusion (blood supply) to the muscles and phonatory mucosa. These data are also critical for improving our ability to use and validate computer models of vocal fold phonatory function that will provide

better insight into the underlying biomechanics associated with normal and pathological voice production, and ultimately help guide the design of improved prevention, diagnostic, and treatment approaches.

4.1. Intraglottal and Vocal Fold Collision Pressure Waveform Characteristics

To the authors' knowledge, the results of this study represented the first time that direct measurement of in vivo vocal fold intraglottal/collision pressures resulted in expected waveform characteristics in a human speaker. In previous work [8], it was found to be more challenging to place a sensor between two vibrating vocal folds in a vocally healthy individual than it was to perform this procedure in an individual with one functional vocal fold (following a hemilaryngectomy surgical procedure). This previous proof-of-concept study demonstrated that it was possible to safely acquire estimates of vocal fold collision forces in one vocally healthy individual (with two functional vocal folds) and three individuals with a hemilaryngectomy. Conclusions from that work were that: (1) simultaneous measurement of subglottal pressure is necessary to accurately interpret the vocal fold collision data, and (2) measures for the individuals with a hemilaryngectomy were easier to obtain and were more reliable and valid than those obtained for the vocally healthy speaker. However, the intraglottal force sensor used in that study had a relatively large sensing area and yielded triangular waveforms that did not appear to exhibit the signature impulse-like component at the time instants of vocal fold collision (an electroglottography signal provided verification of the instants of glottal closure).

Building on this previous work, in the current study, verification of the correct placement of the intraglottal pressure sensor was accomplished using novel, concurrent recordings of laryngeal high-speed videoendoscopy and a subglottal pressure signal. Our prior investigations of a hemilarynx model took advantage of a dual high-speed videoendoscopy setup that offered visual verification of sensor placement from a top-down, endoscopic view and a medial, en face visualization of the of the hemilarynx setup [13]. That work concluded with the following guidelines for using the ISP probe in vivo:

1. Endoscopic visualization is necessary to guide placement of the ISP probe such that the distal pressure sensor at the probe tip is positioned subglottally and the proximal sensor is positioned in the glottis in the phonatory strike zone to sense vocal fold impact collision pressure during phonation.
2. In individuals with a hemilaryngectomy, the ISP probe should rest on the medialized scar band that replaces the excised vocal fold tissue, such that the pressure-sensing element comes into direct contact with the functioning vocal fold.
3. The positioning of the intraglottal pressure sensor is in the phonatory strike zone if the following waveform characteristics are exhibited:
    a. An impulsive peak in the direction of increasing pressure at the instant of vocal fold contact;
    b. A more rounded peak following the impulsive peak that senses aerodynamic pressure build-up during the open phase; and
    c. A minimum value approaching zero or negative pressure immediately preceding the impulsive peak of the subsequent phonatory cycle, reflecting rapidly decreasing intraglottal pressure as airflow accelerates.

Figure 7 displays exemplary intraglottal pressure waveforms observed in the literature according to numerical, physical, excised, and in vivo animal models of phonation. The expected waveform characteristics were observed during the recording in this study as shown in Figure 4.

4.2. Validation of the Two Model-Based Estimates of Vocal Fold Collision Pressure

Although using the ISP probe is not practical for routine data collection, the data collected from the direct measurement of intraglottal, collision, and subglottal pressures using the ISP probe in a select group of individuals can be used as references in numerical models that can be optimized to estimate these important parameters in patients with voice

disorders or vocally healthy speakers. Two numerical modeling approaches were tested in this paper to validate model-based estimates of vocal fold collision pressure using only data from the laryngeal high-speed videoendoscopy recording, a procedure that is feasible to perform in practice. Good agreement was exhibited between the model-estimated collision pressures and measured vocal fold collision pressures. This agreement provides valuable initial in vivo validation of the vocal fold collision pressure estimates derived by these types of models, which is critical, as these engineering methods are expected to be more easily translated into clinical practice.
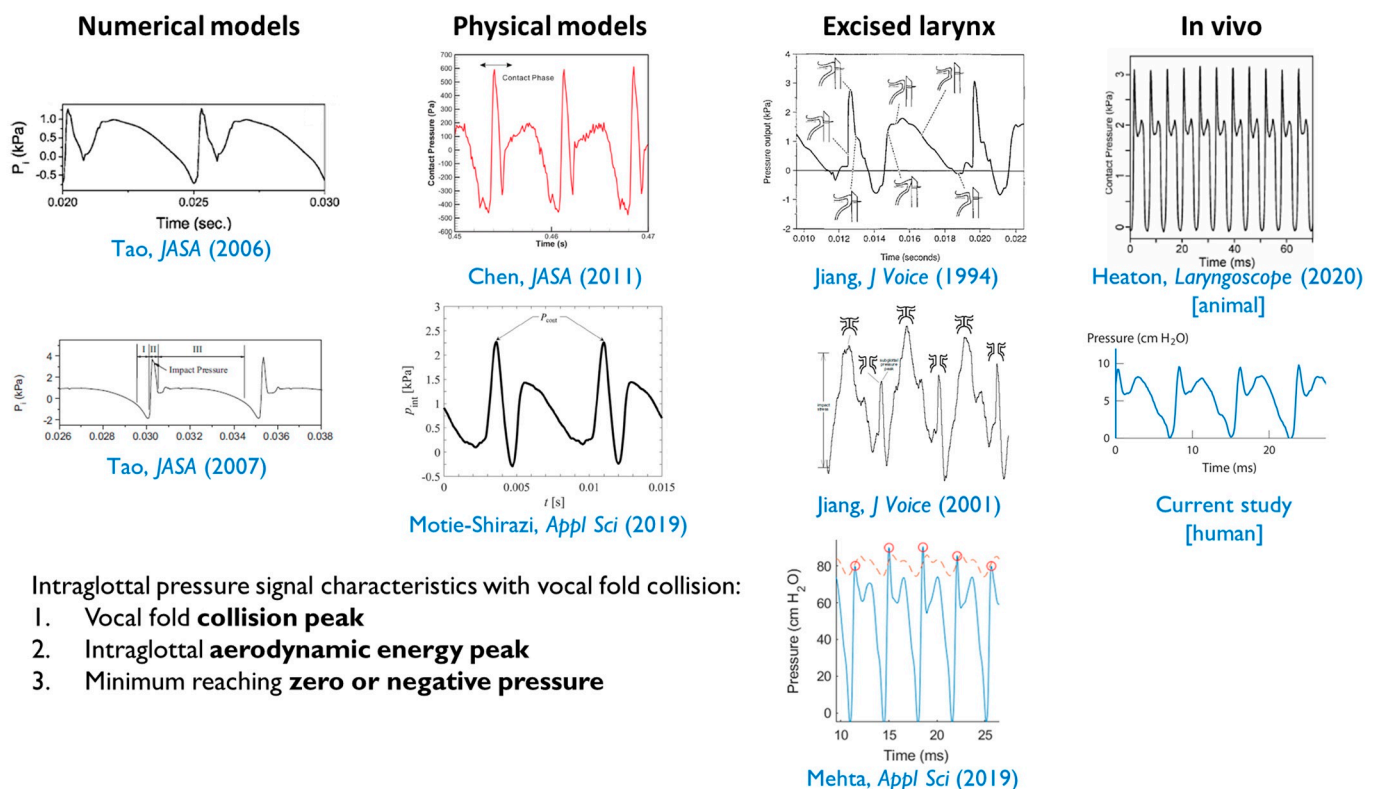
Figure 7. Comparison of the intraglottal pressure signal during phonation of the current study with exemplary intraglottal pressure signals observed in the literature according to numerical models [18,19], physical (synthetic material) models [12,14], excised larynx studies [13,22,25], and an in vivo animal study [29]. Figures reused with permission.

To avoid confirmation bias and given the pilot nature of this study, further in vivo validation is needed to assess the robustness of the indirect estimation approaches. However, prior validation against silicone vocal fold models and prior studies has been successfully performed for both modeling approaches utilized in this study [45,46]. It is interesting to note that both numerical and experimental findings point to a close relation between the peak contact pressure and the driving subglottal pressure, near a 1:1 ratio. The assessment of this relation is of clinical relevance and requires further attention, as other factors are hypothesized to influence this relation, such as muscle activation level, loudness condition, mode of vibration, and supraglottal compression.

4.3. Clinical Implications

One of the most common causes of voice disorders is phonation-related trauma to vocal fold tissue which can cause the loss (e.g., scarring) of normal superficial lamina propria and/or the formation of benign lesions on the vocal folds (e.g., vocal fold nodules). Such damage to the phonatory mucosa is believed to be associated with excessive perilaryngeal muscle activity, termed vocal hyperfunction; however, the actual underlying mechanisms that produce varied vocal fold traumatic injury are poorly understood. The direct measurement of vocal fold collision pressures during phonation, in addition to frictional shearing

stresses on vocal fold tissue and dissipated energy dose [50,51], has important implications in understanding the role of vocal fold collision in the etiology and pathophysiology of phonotraumatic vocal hyperfunction [3]. It is believed that a primary contributing factor to phonotrauma is an increase in level and/or duration above safe thresholds of the collision forces that are generated by the vibrating vocal folds during voice production. However, valid measures of vocal fold collision forces are lacking/incomplete, particularly with respect to what constitutes safe versus damaging levels of such forces.

Of particular interest is the development of measures that can be applied to ambulatory voice monitoring technologies, with the ultimate goal to better differentiate what constitutes healthy versus damaging vocal function. Real-world voice monitoring devices often use sensors placed on the surface of the neck below the larynx as a way to track features of voice production in a way that is confidential, non-obtrusive, and robust to environmental noise artifacts [40,41,52,53]. The vibration accelerometer used in the current study was similarly placed on the neck surface above the sternal notch and below the thyroid prominence to mimic the sensor positioning of an accelerometer used for ambulatory voice monitoring. Several recent studies have started to elucidate vocal features and behaviors during daily life that may be associated with the presence of hyperfunctional voice disorders [39,54–58]. Traditional in-field accelerometer features include estimates of sound pressure level [59] and fundamental frequency that have been combined with voice activity detection to yield vocal dose measures that are designed to indirectly quantify the accumulated effects of phonotrauma [60]. Later formulations of vocal dose measures have attempted to incorporate the effects of vocal fold collision [50], which is important since phonotrauma is hypothesized to be caused by repetitive stress on the mid-membranous portion of the vocal folds [5–7]. Since the data in this study point toward a potential 1:1 correspondence between vocal fold collision pressure and subglottal pressure in certain scenarios, we can take advantage of work aimed at estimating subglottal pressure from neck-surface vibration and apply this analysis to ambulatory voice signals [61–64].

4.4. Study Limitations

It is acknowledged that results should be tempered by the limited analysis of data from a single individual. Given the unique nature of the recording setup, the fact that vocal fold collision characteristics were captured in the intraglottal pressure signal was encouraging. Introducing a probe into the glottis has the potential to disrupt the natural oscillatory mechanisms of the vocal folds. Figure 3 shows the size of the tip of the ISP probe in relation to the size of the glottis and surrounding vocal fold tissue; anterior and posterior glottal gaps may have been introduced due to the presence of the ISP probe. However, despite the introduction of the ISP probe, the study participant was still able to sustain a vowel during the endoscopic procedure and produce relatively stable phonation. Future work is needed to study the effects of varying degrees of loudness levels, pitch conditions, and glottal configurations leading to breathy, modal, and pressed phonation.

Previous work with bench-top physical models of the vocal folds have suggested that caution should be used when interpreting waveform signatures of the intraglottal pressure sensor signal [12]. There could be some uncertainty with respect to whether the intraglottal sensor waveform in Figure 4 was actually from the vocal fold coming into contact with the sensor, or whether the sensor was slightly superior to the strike zone and actually sensing acoustic pressures. In addition, the pressure sensors exhibited signal drifts in the DC level that were also observed in previous studies [8,9]. We believe this drift is primarily caused by thermal energy fluctuations in the glottis due to internal body temperature and the use of a bright xenon light source necessary for high-speed videoendoscopic imaging. Future work could address thermal energy changes by pre-heating the tip of the ISP probe in hot water to match the internal temperature of the glottis.

5. Conclusions

The goal of this study was to illustrate feasibility of the direct measurement of in vivo vocal fold collision pressures during phonation using the recently developed dual-channel ISP probe. The results successfully demonstrated this feasibility in an individual with a hemilaryngectomy, motivating ongoing data collection that is designed to aid in the development of vocal dose measures that incorporate vocal fold impact collision/stress. Even though further in vivo validation with more subjects is needed, the good agreement in this case study between the vocal fold modeling methods and the experimental results illustrates the potential of the modeling tools for both providing access to additional measures of vocal function that are difficult to obtain and advancing the understanding of the underlying control mechanisms of normal and pathological voice production.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/app11167256/s1, Video S1: Laryngeal videostroboscopy of the study participant with a right hemilaryngectomy during sustained phonation at a comfortable pitch and higher pitch. Video S2: Laryngeal high-speed videoendoscopy of vocal fold vibration (2000 frames per second) for the 100 ms phonatory segment during which the intraglottal pressure sensor was in the strike zone. Video S3: Graphical user interface displaying the high-speed video of vocal fold vibration (2000 frames per second) and time-aligned signals (8000 Hz sample rate) recorded by an acoustic microphone, electroglottograph, neck-surface vibration accelerometer, intraglottal pressure, and subglottal pressure. Video S4: Vocal fold edge contour displayed in the high-speed videoendoscopy data for the Hertzian contact pressure model, along with the input and output variables of the model. Signals S5–S9: The time-aligned signal data in WAV format for visualizing the signal waveforms during the 100 ms phonatory segment when vocal fold collision was detected by the intraglottal pressure signal: neck-surface vibration acceleration signal (ACC; S5), electroglottography signal (EGG; S6), ISP probe intraglottal pressure sensor signal (IGP; S7) capturing vocal fold collision pressures during the phonatory closed phase and aerodynamic intraglottal pressure during the phonatory open phase, acoustic microphone signal (MIC; S8), and subglottal pressure signal (SGP; S9).

Author Contributions: Conceptualization, D.D.M. and R.E.H.; data curation, D.D.M.; formal analysis, D.D.M., E.J.I., R.M., and G.A.A.; funding acquisition, D.D.M., M.Z., and R.E.H.; investigation, D.D.M., J.B.K., S.M.Z., R.H.P., and R.E.H.; methodology, D.D.M., J.B.K., S.M.Z., M.Z., B.D.E., S.D.P., and R.E.H.; project administration, R.E.H.; resources, J.B.K., S.M.Z., and R.E.H.; software, D.D.M., E.J.I., G.A.A., and R.M.; supervision, D.D.M. and M.Z.; validation, D.D.M.; visualization, D.D.M., M.Z., E.J.I., G.A.A., and R.M.; writing—original draft preparation, D.D.M. and M.Z.; writing—review and editing, D.D.M., M.Z., B.D.E., S.D.P., and R.E.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Mass General Brigham (protocol number 2008P000652, annual approval obtained on 25 February 2021).

Informed Consent Statement: Informed consent was obtained from the subject involved in the study.

Data Availability Statement: The data presented in this study are available in the Supplementary Materials section.

Conflicts of Interest: Robert Hillman, Steven Zeitels, and Daryush Mehta have a financial interest in InnoVoyce LLC, a company focused on developing and commercializing technologies for the prevention, diagnosis, and treatment of voice-related disorders. Hillman's, Zeitels', and Mehta's interests were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies. Matías Zañartu has a financial interest in Lanek SPA, a company focused on developing and commercializing biomedical devices and technologies. Zañartu's interests were reviewed and are managed by Universidad Técnica Federico Santa María in accordance with its conflict-of-interest policies. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Roy, N.; Merrill, R.M.; Gray, S.D.; Smith, E.M. Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *Laryngoscope* **2005**, *115*, 1988–1995. [CrossRef]
2. Bhattacharyya, N. The prevalence of voice problems among adults in the United States. *Laryngoscope* **2014**, *124*, 2359–2362. [CrossRef]
3. Hillman, R.E.; Stepp, C.E.; Van Stan, J.H.; Zañartu, M.; Mehta, D.D. An updated theoretical framework for vocal hyperfunction. *Am. J. Speech Lang. Pathol.* **2020**, *29*, 2254–2260. [CrossRef]
4. Titze, I.R. Mechanical stress in phonation. *J. Voice* **1994**, *8*, 99–105. [CrossRef]
5. Czerwonka, L.; Jiang, J.J.; Tao, C. Vocal nodules and edema may be due to vibration-induced rises in capillary pressure. *Laryngoscope* **2008**, *118*, 748–752. [CrossRef]
6. Tao, C.; Jiang, J.J.; Czerwonka, L. Liquid accumulation in vibrating vocal fold tissue: A simplified model based on a fluid-saturated porous solid theory. *J. Voice* **2010**, *24*, 260–269. [CrossRef] [PubMed]
7. Kvit, A.A.; Devine, E.E.; Jiang, J.J.; Vamos, A.C.; Tao, C. Characterizing liquid redistribution in a biphasic vibrating vocal fold using finite element analysis. *J. Voice* **2015**, *29*, 265–272. [CrossRef]
8. Gunter, H.E.; Howe, R.D.; Zeitels, S.M.; Kobler, J.B.; Hillman, R.E. Measurement of vocal fold collision forces during phonation: Methods and preliminary data. *J. Speech Lang. Hear. Res.* **2005**, *48*, 567–576. [CrossRef]
9. Verdolini, K.; Hess, M.M.; Titze, I.R.; Bierhals, W.; Gross, M. Investigation of vocal fold impact stress in human subjects. *J. Voice* **1999**, *13*, 184–202. [CrossRef]
10. Hess, M.M.; Verdolini, K.; Bierhals, W.; Mansmann, U.; Gross, M. Endolaryngeal contact pressures. *J. Voice* **1998**, *12*, 50–67. [CrossRef]
11. Rothenberg, M. A multichannel electroglottograph. *J. Voice* **1992**, *6*, 36–43. [CrossRef]
12. Motie-Shirazi, M.; Zañartu, M.; Peterson, S.D.; Mehta, D.D.; Kobler, J.B.; Hillman, R.E.; Erath, B.D. Toward development of a vocal fold contact pressure probe: Sensor characterization and validation using synthetic vocal fold models. *Appl. Sci.* **2019**, *9*, 3002. [CrossRef]
13. Mehta, D.D.; Kobler, J.B.; Zeitels, S.M.; Zañartu, M.; Erath, B.D.; Motie-Shirazi, M.; Peterson, S.D.; Petrillo, R.H.; Hillman, R.E. Toward development of a vocal fold contact pressure probe: Bench-top validation of a dual-sensor probe using excised human larynx models. *Appl. Sci.* **2019**, *9*, 4360. [CrossRef] [PubMed]
14. Chen, L.-J.; Mongeau, L. Verification of two minimally invasive methods for the estimation of the contact pressure in human vocal folds during phonation. *J. Acoust. Soc. Am.* **2011**, *130*, 1618–1627. [CrossRef] [PubMed]
15. Gunter, H.E. A mechanical model of vocal-fold collision with high spatial and temporal resolution. *J. Acoust. Soc. Am.* **2003**, *113*, 994–1000. [CrossRef]
16. Gunter, H.E. Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions. *J. Biomech.* **2004**, *37*, 1119–1124. [CrossRef]
17. Horáček, J.; Šidlof, P.; Švec, J.G. Numerical simulation of self-oscillations of human vocal folds with Hertz model of impact forces. *J. Fluids Struct.* **2005**, *20*, 853–869. [CrossRef]
18. Tao, C.; Jiang, J.J.; Zhang, Y. Simulation of vocal fold impact pressures with a self-oscillating finite-element model. *J. Acoust. Soc. Am.* **2006**, *119*, 3987–3994. [CrossRef]
19. Tao, C.; Jiang, J.J. Mechanical stress during phonation in a self-oscillating finite-element vocal fold model. *J. Biomech.* **2007**, *40*, 2191–2198. [CrossRef]
20. Horáček, J.; Laukkanen, A.M.; Šidlof, P.; Murphy, P.; Švec, J.G. Comparison of acceleration and impact stress as possible loading factors in phonation: A computer modeling study. *Folia Phoniatr. Logop.* **2009**, *61*, 137–145. [CrossRef] [PubMed]
21. Spencer, M.; Siegmund, T.; Mongeau, L. Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation. *J. Acoust. Soc. Am.* **2008**, *123*, 1089–1103. [CrossRef] [PubMed]
22. Jiang, J.J.; Titze, I.R. Measurement of vocal fold intraglottal pressure and impact stress. *J. Voice* **1994**, *8*, 132–144. [CrossRef]
23. Verdolini, K.; Chan, R.; Titze, I.R.; Hess, M.; Bierhals, W. Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges. *J. Voice* **1998**, *12*, 415–423. [CrossRef]
24. Berry, D.A.; Verdolini, K.; Montequin, D.W.; Hess, M.M.; Chan, R.W.; Titze, I.R. A quantitative output-cost ratio in voice production. *J. Speech Lang. Hear. Res.* **2001**, *44*, 29–37. [CrossRef]

25. Jiang, J.J.; Shah, A.G.; Hess, M.M.; Verdolini, K.; Banzali, F.M., Jr.; Hanson, D.G. Vocal fold impact stress analysis. *J. Voice* **2001**, *15*, 4–14. [CrossRef]

26. Heaton, J.T.; Kobler, J.B.; Hillman, R.E.; Zeitels, S.M. A new instrument for intraoperative assessment of individual vocal folds. *Laryngoscope* **2005**, *115*, 1223–1229. [CrossRef]

27. Backshaei, H.; Yang, J.; Miri, A.K.; Mongeau, L. Determination of the stresses and strain on the superior surface of excised porcine larynges during phonation using digital image correlation. *Proc. Meet. Acoust.* **2013**, *19*, 060238.

28. Weiss, S.; Sutor, A.; Rupitsch, S.J.; Kniesburges, S.; Doellinger, M.; Lerch, R. Development of a small film sensor for the estimation of the contact pressure of artificial vocal folds. *Proc. Meet. Acoust.* **2013**, *19*, 060307.

29. Heaton, J.T.; Kobler, J.B.; Ottensmeyer, M.P.; Petrillo, R.H.; Tynan, M.A.; Mehta, D.D.; Hillman, R.E.; Zeitels, S.M. Aerodynamically driven phonation of individual vocal folds under general anesthesia in canines. *Laryngoscope* **2020**, *130*, 1980–1988. [CrossRef]

30. Zeitels, S.M.; Jarboe, J.; Franco, R.A. Phonosurgical reconstruction of early glottic cancer. *Laryngoscope* **2001**, *111*, 1862–1865. [CrossRef]

31. Zeitels, S.M.; Hillman, R.E.; Franco, R.A.; Bunting, G.W. Voice and treatment outcome from phonosurgical management of early glottic cancer. *Ann. Otol. Rhinol. Laryngol.* **2002**, *111*, 1–20. [CrossRef]

32. Zeitels, S.M. Optimizing voice after endoscopic partial laryngectomy. *Otolaryngol. Clin. N. Am.* **2004**, *37*, 627–636. [CrossRef]

33. Alipour, F.; Montequin, D.; Tayama, N. Aerodynamic profiles of a hemilarynx with a vocal tract. *Ann. Otol. Rhinol. Laryngol.* **2001**, *110*, 550–555. [CrossRef] [PubMed]

34. Doellinger, M.; Berry, D.A. Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation. *J. Voice* **2006**, *20*, 401–413. [CrossRef] [PubMed]

35. Jiang, J.J.; Titze, I.R. A methodological study of hemilaryngeal phonation. *Laryngoscope* **1993**, *103*, 872–882. [CrossRef]

36. Saraniti, C.; Speciale, R.; Santangelo, M.; Massaro, N.; Maniaci, A.; Gallina, S.; Serra, A.; Cocuzza, S. Functional outcomes after supracricoid modified partial laryngectomy. *J. Biol. Regul. Homeost. Agents* **2019**, *33*, 1903–1907. [PubMed]

37. Zañartu, M.; Ho, J.C.; Mehta, D.D.; Hillman, R.E.; Wodicka, G.R. Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Trans. Audio Speech Lang. Process.* **2013**, *21*, 1929–1939. [CrossRef]

38. Cheyne, H.A. Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck. In Proceedings of the 3rd IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors, Cambridge, MA, USA, 4–6 September 2006; pp. 118–121.

39. Cortés, J.P.; Espinoza, V.M.; Ghassemi, M.; Mehta, D.D.; Van Stan, J.H.; Hillman, R.E.; Guttag, J.V.; Zañartu, M. Ambulatory assessment of phonotraumatic vocal hyperfunction using glottal airflow measures estimated from neck-surface acceleration. *PLoS ONE* **2018**, *13*, e0209017. [CrossRef] [PubMed]

40. Mehta, D.D.; Van Stan, J.H.; Zañartu, M.; Ghassemi, M.; Guttag, J.V.; Espinoza, V.M.; Cortés, J.P.; Cheyne, H.A., II; Hillman, R.E. Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Front. Bioeng. Biotechnol.* **2015**, *3*, 1–14. [CrossRef] [PubMed]

41. Popolo, P.S.; Švec, J.G.; Titze, I.R. Adaptation of a Pocket PC for use as a wearable voice dosimeter. *J. Speech Lang. Hear. Res.* **2005**, *48*, 780–791. [CrossRef]

42. Cheyne, H.A.; Hanson, H.M.; Genereux, R.P.; Stevens, K.N.; Hillman, R.E. Development and testing of a portable vocal accumulator. *J. Speech Lang. Hear. Res.* **2003**, *46*, 1457–1467. [CrossRef]

43. Winholtz, W.S.; Titze, I.R. Conversion of a head-mounted microphone signal into calibrated SPL units. *J. Voice* **1997**, *11*, 417–421. [CrossRef]

44. Boersma, P.; Weenink, D. *Praat: Doing Phonetics by Computer*; University of Amsterdam: Amsterdam, The Netherlands, 2003; Available online: http://www.praat.org (accessed on 21 July 2003).

45. Díaz-Cádiz, M.E.; Peterson, S.D.; Galindo, G.E.; Espinoza, V.M.; Motie-Shirazi, M.; Erath, B.D.; Zañartu, M. Estimating vocal fold contact pressure from raw laryngeal high-speed videoendoscopy using a Hertz contact model. *Appl. Sci.* **2019**, *9*, 2384. [CrossRef]

46. Alzamendi, G.A.; Manríquez, R.; Hadwin, P.J.; Deng, J.J.; Peterson, S.D.; Erath, B.D.; Mehta, D.D.; Hillman, R.E.; Zañartu, M. Bayesian estimation of vocal function measures using laryngeal high-speed videoendoscopy and glottal airflow estimates: An in vivo case study. *J. Acoust. Soc. Am.* **2020**, *147*, EL434–EL439. [CrossRef] [PubMed]

47. Zañartu, M.; Galindo, G.E.; Erath, B.D.; Peterson, S.D.; Wodicka, G.R.; Hillman, R.E. Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *J. Acoust. Soc. Am.* **2014**, *136*, 3262–3271. [CrossRef] [PubMed]

48. Birkholz, P. GlottalImageExplorer—An open source tool for glottis segmentation in endoscopic high-speed videos of the vocal folds. In *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung*; Jokisch, O., Ed.; TUDPress: Dresden, Germany, 2016.

49. Titze, I.R. Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *J. Speech Lang. Hear. Res.* **2006**, *49*, 439–447. [CrossRef]

50. Titze, I.R.; Hunter, E.J. Comparison of vocal vibration-dose measures for potential-damage risk criteria. *J. Speech Lang. Hear. Res.* **2015**, *58*, 1425–1439. [CrossRef]

51. Motie-Shirazi, M.; Zañartu, M.; Peterson, S.D.; Erath, B.D. Vocal fold dynamics in a synthetic self-oscillating model: Contact pressure and dissipated energy dose. *J. Acoust. Soc. Am.* **2021**, *150*, 478–489. [CrossRef]

52. Hillman, R.E.; Heaton, J.T.; Masaki, A.; Zeitels, S.M.; Cheyne, H.A. Ambulatory monitoring of disordered voices. *Ann. Otol. Rhinol. Laryngol.* **2006**, *115*, 795–801. [CrossRef]
53. Szabo Portela, A.; Granqvist, S.; Ternström, S.; Södersten, M. Vocal behavior in environmental noise: Comparisons between work and leisure conditions in women with work-related voice disorders and matched controls. *J. Voice* **2018**, *32*, 126.e23–126.e38. [CrossRef]
54. Van Stan, J.H.; Mehta, D.D.; Ortiz, A.J.; Burns, J.A.; Toles, L.E.; Marks, K.L.; Vangel, M.; Hron, T.; Zeitels, S.; Hillman, R.E. Differences in weeklong ambulatory vocal behavior between female patients with phonotraumatic lesions and matched controls. *J. Speech Lang. Hear. Res.* **2020**, *63*, 372–384. [CrossRef]
55. Van Stan, J.H.; Mehta, D.D.; Ortiz, A.J.; Burns, J.A.; Marks, K.L.; Toles, L.E.; Stadelman-Cohen, T.; Krusemark, C.; Muise, J.; Hron, T.; et al. Changes in a Daily Phonotrauma Index after laryngeal surgery and voice therapy: Implications for the role of daily voice use in the etiology and pathophysiology of phonotraumatic vocal hyperfunction. *J. Speech Lang. Hear. Res.* **2020**, *63*, 3934–3944. [CrossRef] [PubMed]
56. Van Stan, J.H.; Ortiz, A.J.; Cortés, J.P.; Marks, K.L.; Toles, L.E.; Mehta, D.D.; Burns, J.A.; Hron, T.; Stadelman-Cohen, T.; Krusemark, C.; et al. Differences in daily voice use measures between female patients with nonphonotraumatic vocal hyperfunction and matched controls. *J. Speech Lang. Hear. Res.* **2021**, *64*, 1457–1470. [CrossRef] [PubMed]
57. Van Stan, J.; Ortiz, A.; Marks, K.; Toles, L.; Mehta, D.; Burns, J.; Hron, T.; Stadelman-Cohen, T.; Krusemark, C.; Muise, J.; et al. Changes in the Daily Phonotrauma Index (DPI) following the use of voice therapy as the sole treatment for phonotraumatic vocal hyperfunction in females. *J. Speech Lang. Hear. Res.* **2021**, in press.
58. Toles, L.E.; Ortiz, A.J.; Marks, K.L.; Burns, J.A.; Hron, T.; Van Stan, J.H.; Mehta, D.D.; Hillman, R.E. Differences between female singers with phonotrauma and vocally healthy matched controls in singing and speaking voice use during 1 week of ambulatory monitoring. *Am. J. Speech Lang. Pathol.* **2021**, *30*, 199–209. [CrossRef] [PubMed]
59. Švec, J.G.; Titze, I.R.; Popolo, P.S. Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *J. Acoust. Soc. Am.* **2005**, *117*, 1386–1394. [CrossRef] [PubMed]
60. Titze, I.R.; Švec, J.G.; Popolo, P.S. Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *J. Speech Lang. Hear. Res.* **2003**, *46*, 919–932. [CrossRef]
61. Fryd, A.S.; Van Stan, J.H.; Hillman, R.E.; Mehta, D.D. Estimating subglottal pressure from neck-surface acceleration during normal voice production. *J. Speech Lang. Hear. Res.* **2016**, *59*, 1335–1345. [CrossRef]
62. Marks, K.L.; Lin, J.Z.; Fox, A.B.; Toles, L.E.; Mehta, D.D. Impact of nonmodal phonation on estimates of subglottal pressure from neck-surface acceleration in healthy speakers. *J. Speech Lang. Hear. Res.* **2019**, *62*, 3339–3358. [CrossRef]
63. Lin, J.Z.; Espinoza, V.M.; Marks, K.L.; Zañartu, M.; Mehta, D.D. Improved subglottal pressure estimation from neck-surface vibration in healthy speakers producing non-modal phonation. *IEEE J. Sel. Top. Signal Process.* **2020**, *14*, 449–460. [CrossRef]
64. Marks, K.; Lin, J.Z.; Burns, J.A.; Hron, T.A.; Hillman, R.E.; Mehta, D.D. Estimation of subglottal pressure from neck surface vibration in patients with voice disorders. *J. Speech Lang. Hear. Res.* **2020**, *63*, 2202–2218. [CrossRef]