JSLHR

Research Note

Estimating Subglottal Pressure From Neck-Surface Acceleration During Normal Voice Production

Amanda S. Fryd,[a,b] Jarrad H. Van Stan,[a,b] Robert E. Hillman,[a,b,c,d] and Daryush D. Mehta[a,b,c]

Purpose: The purpose of this study was to evaluate the potential for estimating subglottal air pressure using a neck-surface accelerometer and to compare the accuracy of predicting subglottal air pressure relative to predicting acoustic sound pressure level (SPL).
Method: Indirect estimates of subglottal pressure (P_{sg}') were obtained from 10 vocally healthy speakers during loud-to-soft repetitions of 3 different /p/–vowel gestures (/pa/, /pi/, /pu/) at 3 pitch levels in the modal register. Intraoral air pressure, neck-surface acceleration, and radiated acoustic pressure were recorded, and the root-mean-square amplitude of the acceleration signal was correlated with P_{sg}' and SPL.
Results: The coefficient of determination between accelerometer level and P_{sg}' was high when data were pooled from all vowel and pitch contexts for each participant (r^2 = .68–.93). These relationships were stronger than corresponding relationships between accelerometer level and SPL (r^2 = .46–.81). The average 95% prediction interval for estimating P_{sg}' using accelerometer level was ±2.53 cm H_2O, ranging from ±1.70 to ±3.74 cm H_2O across participants.
Conclusions: Accelerometer signal amplitude correlated more strongly with P_{sg}' than with SPL. Future work is warranted to investigate the robustness of the relationship in nonmodal voice qualities, individuals with voice disorders, and accelerometer-based ambulatory monitoring of subglottal pressure.

Voice specialists make critical diagnostic, medical, therapeutic, and surgical decisions on the basis of coupling visual observations of vocal-fold tissue motion with auditory-perceptual assessments of voice quality (Zeitels, Blitzer, Hillman, & Anderson, 2007). Objective acoustic, aerodynamic, and electroglottographic measures are often used to better document the impact of disorders on vocal function and to help assess the effects of treatment (Hillman, Montgomery, & Zeitels, 1997; Roy et al., 2013). There is ongoing interest in obtaining measures of vocal function through the use of a small accelerometer (ACC) placed on the anterior neck surface below the larynx (subglottal) to record neck-surface vibrations as a user phonates, including estimates of aerodynamic parameters (Mehta, Van Stan, & Hillman, 2016; Wokurek & Pützer, 2009; Zañartu, Ho, Mehta, Hillman, & Wodicka, 2013). During phonation, the subglottal ACC waveform is primarily influenced by acoustic and aerodynamic energy radiating from the trachea and is minimally affected by supraglottal resonances (Wokurek & Madsack, 2011; Zañartu, Ho, et al., 2013). This presents the possibility that more robust estimates of some voicing source parameters could be extracted directly from the ACC signal than can be obtained by trying to correlate ACC signal properties with vocal parameters extracted from an acoustic microphone (MIC) recording, because of the variable influence of supraglottal resonances on the MIC signal.

To be specific, the average uncertainty of predicting acoustic sound pressure level (SPL) using the root-mean-square amplitude of the ACC signal has been reported to be ±6 and ±5 dB for male and female adult speakers, respectively (Švec, Titze, & Popolo, 2005). This degree of uncertainty can be problematic, because SPL estimates obtained from ACC data are often used to derive higher-level voice-use parameters, such as long-term vocal-dose measures (Titze, Švec, & Popolo, 2003). Vocal-dose measures have been thought to be clinically important and are applied to the assessment of individuals in high-voice-use

[a]Communication Sciences and Disorders, MGH Institute of Health Professions, Charlestown, MA
[b]Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston
[c]Department of Surgery, Harvard Medical School, Boston, MA
[d]Surgery & Health Sciences and Technology, Harvard Medical School, Boston, MA

Correspondence to Daryush D. Mehta: daryush.mehta@alum.mit.edu

Editor: Julie Liss
Associate Editor: Nelson Roy
Received December 14, 2015
Revision received March 10, 2016
Accepted May 6, 2016
DOI: 10.1044/2016_JSLHR-S-15-0430

Disclosure: The authors have declared that no competing interests existed at the time of publication.

occupations (Bottalico & Astolfi, 2012; Lindstrom, Waye, Södersten, McAllister, & Ternström, 2011; Titze & Hunter, 2015) as well as individuals diagnosed with behaviorally based voice disorders (Mehta et al., 2015; Nacci et al., 2013; Van Stan et al., 2015). Thus, improving the accuracy of ACC-based estimates of vocal parameters would have a positive impact on voice research and ultimately on the application of this technology to enhance clinical practice.

Recent studies have supported further investigation into estimating aerodynamic voice measurements from neck-surface acceleration signals. Zañartu, Ho, et al. (2013) validated estimates of the glottal-airflow waveform derived from the ACC signal, with inverse-filtered oral airflow acting as the reference signal. ACC signals were processed using subglottal impedance-based inverse filtering, a model-based filtering technique designed to extract the glottal-airflow waveform and associated waveform properties. Results indicated that ACC-derived values were comparable to those gained through oral-airflow analysis of sustained vowel production. To be specific, two glottal-airflow features extracted during the open phase of the glottal cycle (peak-to-peak airflow and maximum flow declination rate) were strongly correlated with features derived from the reference oral-airflow recordings ($r^2 = .98$ and .97, respectively). Categorization of individuals exhibiting vocal hyperfunction versus normal phonation is beginning to show promise using glottal-airflow features (Mehta et al., 2015; Zañartu, Espinoza et al., 2013).

One relationship yet to be investigated is that between ACC signal properties and average subglottal air pressure, leading to the transglottal pressure that drives the vocal folds into self-sustained oscillation during phonation. Subglottal pressure is a key factor in voice production, because it contributes to the vibratory behavior of the larynx (Holmberg, Hillman, & Perkell, 1988; Stevens, 1998) and has been shown to directly affect both vocal intensity (Holmberg et al., 1988; Tanaka & Gould, 1983) and fundamental frequency (Holmberg, Hillman, & Perkell, 1989; Titze, 1989) in human subjects, as well as in excised-larynx experiments (for a review, see Döllinger et al., 2011). Related work has also investigated empirical relationships between vocal-fold vibratory amplitude and anterior neck-surface acceleration during phonation in human subjects (Popolo, 2007).

It is hypothesized that the subglottal ACC-signal amplitude will result in a stronger and less variable correlation with subglottal-pressure estimates (P_{sg}') than with acoustic SPL across different vowels and fundamental frequencies. The inherent SPL of running speech is significantly influenced by vocal-tract resonances that change dynamically in time. This may be the main factor resulting in the imprecise estimates of SPL derived from ACC-signal amplitude (Švec et al., 2005). In contrast, tracking subglottal pressure using the ACC signal may result in reduced uncertainty due to the ACC signal's being minimally affected by supraglottal resonances across different vowel contexts (Cheyne, Hanson, Genereux, Stevens, & Hillman, 2003; Zañartu, Ho, et al., 2013).

Clinical Importance

Subglottal pressure has been used clinically to differentiate individuals with typical voices from those with voice disorders and to act as a clinical outcome measure (Hartl, Hans, Vaissière, Riquet, & Brasnu, 2001; Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989; Holmberg, Doyle, Perkell, Hammarberg, & Hillman, 2003; Jiang & Stern, 2004; Speyer, 2008; Zeitels, Hillman, Franco, & Bunting, 2002; Zeitels, Hochman, & Hillman, 1998; Zeitels et al., 2009). In addition, subglottal pressure is a central component of vocal efficiency (Björklund & Sundberg, 2016; Colton, Casper, & Leonard, 2006; Titze, 1992, 2013; Titze, Maxfield, & Palaparthi, 2016) and is associated with aspects of perceived vocal effort (Ramig & Dromey, 1996; Rosenthal, Lowell, & Colton, 2014). Empirical evidence from individuals with voice disorders has pointed to elevated levels of subglottal pressure required to produce a particular SPL compared with speakers with typical voices (Hillman et al., 1989; Netsell, Lotz, & Shaughnessy, 1984). For example, the ratio of SPL to P_{sg}' is hypothesized to reflect aspects of vocal effort in individuals developing vocal problems who report that they need to expend more energy than they did before their voice difficulties began (Colton et al., 2006). Furthermore, increased P_{sg}' has been theoretically related to phonotraumatic and/or fatiguing vocal behaviors (Hillman et al., 1989; Zañartu et al., 2014).

Measurements of subglottal pressure, however, are underutilized in clinical settings because current methods of obtaining subglottal pressure do not readily allow for its estimation during running speech, or do so through invasive techniques or sophisticated or expensive equipment (Cranen & Boves, 1985; Hixon, 1972; Plant & Hillel, 1998; Sundberg, Scherer, Hess, Müller, & Granqvist, 2013; Tanaka & Gould, 1983; van den Berg, 1956). In contrast, noninvasive estimation of vocal-function measures such as P_{sg}' can be obtained from an unobtrusive ACC signal during natural speech production in laboratory, clinical, and ambulatory settings. In particular, recent developments have allowed for long-term, ambulatory monitoring of voice use in naturalistic settings by an ACC sensor attached to a smartphone interface (Mehta, Zañartu, Feng, Cheyne, & Hillman, 2012). Monitoring vocal function with an ACC-based method has great potential to improve the assessment and treatment of voice disorders during in-clinic therapy sessions and ambulatory monitoring and/or biofeedback using the patient's actual phonatory characteristics during activities of daily living.

This initial study of individuals with typical voices served as the first step toward identifying and quantifying the relationship between P_{sg}' and subglottal neck-surface vibration. In particular, the study aimed to answer the following questions:

1. What is the correlation between ACC-signal amplitude and P_{sg}' during sustained vowel production?

2. What is the variation in the correlation between P_{sg}' and ACC-signal amplitude across different vowel and fundamental-frequency conditions?
3. Does the relationship between P_{sg}' and ACC-signal amplitude result in less uncertainty than that between the ACC signal and acoustic SPL across vowel contexts and fundamental frequencies?

Method

Participants

Ten healthy adult speakers (five men, five women) were selected from a pool of English-speaking adults who were enrolled in a larger study investigating smartphone-based ambulatory voice monitoring (Mehta et al., 2015). The mean (*SD*) participant age was 22.7 (5.8) years, ranging from 18 to 33 years. These individuals underwent transoral rigid endoscopy by a licensed speech-language pathologist to verify typical vocal status.

Data Collection

In a sound-treated booth, estimates of subglottal air pressure were obtained via intraoral pressure recordings during an occlusive plosive using an intraoral catheter connected to a pressure sensor (Glottal Enterprises Inc., Syracuse, NY). Participants repeated consonant–vowel pairs in modal voice from loud to soft levels within three vowel contexts (/pa/, /pi/, /pu/) and at three pitch levels (comfortable, lower than comfortable, higher than comfortable). The elicitation of /p/–vowel pairs with gradually decreasing loudness allowed for the collection of a larger spectrum of SPL data points than would have been obtained using a standard protocol that calls for a single loudness level per trial (Rothenberg, 1973). Further, this method was time efficient and preferable to avoid participant fatigue, which could have led to an increase in vocal effort over the experimental duration. Björklund and Sundberg (2016) used similar "de/crescendo" protocols to study the relationship between P_{sg}' and SPL, resulting in a high correlation (r = .83) over a wide range of P_{sg}' values. Vocal-intensity changes have also been assessed using semioccluded-vocal-tract methods in a context of varying pitch and vowel (Titze & Hunter, 2011); note, though, that we chose to estimate P_{sg}' during consonant rather than vocalic segments. The three vowel contexts (/a/, /i/, /u/) were chosen on the basis of their approximate cardinal positions on the American English vowel formant space (Hillenbrand, Getty, Clark, & Wheeler, 1995), in order to detect changes in the ACC signal that simulate the variations occurring during continuous speech production.

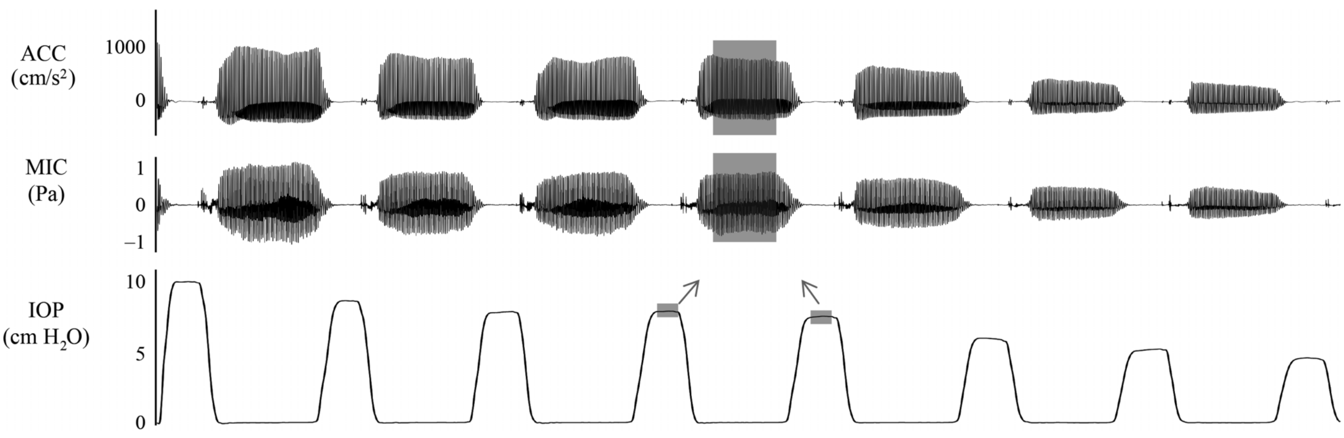
During a training phase, study staff demonstrated the /p/–vowel task and participants were instructed to produce the tokens at their naturally comfortable pitch—that is, not to mimic the pitch of the trainer. Predetermined fundamental frequencies were not requested from participants, out of a desire to keep phonation as similar as possible to each participant's relative habitual voice production (Holmberg et al., 1989). The average fundamental frequencies naturally elicited for (respectively) the comfortable, lower, and higher pitch conditions were, for female participants, 300 Hz (baseline), 235 Hz (baseline − 4.2 semitones [STs]), and 396 Hz (baseline + 4.8 STs); and for male participants, 156 Hz (baseline), 122 Hz (baseline − 4.3 STs), and 202 Hz (baseline + 4.5 STs). Fundamental frequencies elicited within each pitch condition were generally higher than is typical for connected speech, which was largely attributed to the artificial nature of the task (i.e., production of a series of loudness levels from loud to soft during each syllable string) and the natural elevation of fundamental frequency when starting each syllable string at a high SPL; participants were tasked to maintain their pitch for the duration of the syllable string.

Simultaneous recordings were made using an intraoral pressure (IOP) sensor, a high-bandwidth ACC sensor (BU-27135, Knowles Electronics, Itasca, IL), and a head-mounted condenser MIC positioned approximately 10 cm from the participant's lips (MKE104, Sennheiser Electronic GmbH, Wennebostel, Germany). The experimenter affixed the ACC to the participant's neck halfway between the thyroid prominence and the suprasternal notch using hypoallergenic double-sided tape (Model 2181, 3M, Maplewood, MN). The MIC signal was input to a preamplifier (Model 302 Dual Microphone Preamplifier, Symetrix, Inc., Mountlake Terrace, WA). All three signals were routed through a preamplifier that provided low-pass antialiasing filtering with a cutoff frequency of 8 kHz (CyberAmp Model 380, Axon Instruments, Union City, CA) prior to digital sampling at 20 kHz and 16-bit quantization (Digidata 1440A, Axon Instruments).

IOP, ACC, and MIC signals were calibrated to physical units. The IOP transducer was calibrated using a closed syringe system that provided reference levels of 0, 5, 10, 15, and 20 cm H_2O. The MIC signal was calibrated using a Cooper-Rand electrolarynx sound source that generated multiple reference tones at increasing intensity levels measured by a Class 2 sound-level meter (NL-20, Rion, Tokyo, Japan) to map the uncalibrated voltage signal to units of pascals and dB SPL at 10 cm. The ACC assembly (sensor mounted on a silicone pad) was calibrated to physical units of acceleration (cm/s^2) by sending a known stimulus amplitude to a mechanical shaker (Cheyne, 2002).

Figure 1 illustrates signals obtained from one female participant (F1) during one trial of /p/–vowel (/pi/) pairs in loud to soft voice. To estimate subglottal pressure from the IOP signal, the expected buildup of pressure during the /p/ consonant must be followed by a relatively constant pressure signal to indicate that air pressure has equilibrated throughout the airway before the lips open for the subsequent vowel. If the trial was successful, the peaks of the pressure signal appeared as plateaus, and IOP was considered equilibrated with subglottal pressure. Participants were provided verbal instruction and were able to see the computer screen to coach them in accomplishing this task.

Figure 1. Example waveforms of simultaneously recorded signals of a neck-surface accelerometer (ACC), an acoustic microphone (MIC), and intraoral air pressure (IOP) during the production of consonant–vowel (/pi/) pairs with decreasing loudness performed by female participant F1. For ACC and MIC signals, shaded regions represent manually segmented boundaries during a given vowel from which the root-mean-square amplitudes were computed. For the IOP signal, shaded regions indicate manually segmented IOP boundaries used to measure various waveform characteristics. During processing, mean IOP values before and after the vowel were averaged to obtain an estimated subglottal pressure during the vowel.



Data Analysis

In Figure 1, shading in the IOP signal exemplifies the manual segmentation of the boundaries of each plateau during bilabial occlusion using Praat (Boersma & Weenink, 2014). Using custom MATLAB code, IOP waveforms were low-pass filtered (80-Hz cutoff frequency) in order to remove harmonic information from the recorded signal. The mean, standard deviation, minimum, and maximum were computed for each IOP plateau. As seen in Figure 1, P_{sg}' for each vowel production was determined by taking the mean of the average IOP plateaus that preceded and followed a given vowel. The peak-to-peak value of a given pressure plateau was computed by subtracting the minimum from the maximum pressure value. Shading in the ACC and MIC signals exemplifies the manual segmentation of the steady-state portions of each vowel. The root-mean-square amplitudes of the ACC and MIC signals were computed.

To assume full equilibration of air pressure from the lungs to the oral cavity, the IOP signal needed to be as flat as possible (Rothenberg, 1973). However, there is currently no standardized method of evaluating whether an IOP signal is "flat enough" to assume air-pressure equilibration. Previous studies (Löfqvist, Carlborg, & Kitzing, 1982; Plant & Hillel, 1998) accepted all pressure plateaus without any attempt to remove undesirable "peaky" signals that often result from respiratory pumping or excessive aspiration during /p/ production. The frequency content of these peaks was not significantly affected by the low-pass filter (80-Hz cutoff frequency) that was applied to remove harmonic information; the filter preserved undesirable peaks so that they could be identified. Because the goal of this study was to establish a baseline relationship between ACC-signal amplitude and P_{sg}', only trials yielding the most accurate estimates of P_{sg}' were desired. A conservative peak-to-peak threshold (potentially screening out satisfactory plateaus) of 0.5 cm H_2O was visually validated to reflect sufficient plateau stability for further analysis. Thus, a consistent screening process across participants omitted IOP peaks whose peak-to-peak amplitudes were larger than this threshold; the screening yielded an average of 327 vowel segments for each participant for inclusion in the statistical analysis.

Statistical Analysis

Three performance metrics were computed. First, the coefficient of determination (r^2) was used to analyze the relationship between ACC-signal amplitude and P_{sg}' within and across the different vowel and fundamental-frequency contexts, with the aim of comparing the strengths of these correlations with those computed between ACC-signal amplitude and SPL. Linear regressions were performed to compute the variation, slope, and intercept of the mapping from ACC-signal amplitude to both P_{sg}' and SPL, given changes in vowel and fundamental frequency.

Second, average 95% prediction intervals (PI_{95}) were computed for each participant across all pitch and vowel contexts for comparison with published degrees of uncertainty. Recall that Švec et al. (2005) reported an average PI_{95} of ±6 dB for the prediction of SPL using ACC-signal amplitude for running speech passages. Obtaining similar PI_{95} values for ACC-derived SPL in the current study would provide evidence supporting the selection of only three vowels and three pitch levels. In addition, analogous PI_{95} values were computed for ACC-derived P_{sg}'—that is, the prediction of P_{sg}' using ACC-signal amplitude.

The PI_{95} for estimating a specific SPL or P_{sg}' value (y variable) given ACC-signal amplitude (x variable) was calculated using the equation

$$PI_{95} = y \pm t_{crit} \cdot SEP, \quad (1)$$

where y is the predicted value of x given the corresponding linear regression equation across the data points, t_{crit} is the critical value for a two-tailed t distribution (α = .05) with $n - 2$ degrees of freedom, and SEP is the standard error of the prediction:

$$SEP = SEM\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}, \quad (2)$$

where SEM is the standard error of the mean for predicting y given x, and $\bar{x}$ is the mean of all x values. The reported PI_{95} values for each scatter plot were mean values over all predicted data points.

Third, a more direct comparison was desired to determine whether ACC-derived P_{sg}' had more or less uncertainty than ACC-derived SPL because the two uncertainties were in different units—cm H_2O and dB SPL, respectively. To this end, a method of mapping P_{sg}' uncertainty onto SPL uncertainty was devised. SPL has been empirically shown to increase by approximately 13 dB for every doubling of P_{sg}' in adult speakers (Holmberg et al., 1988). Similar relationships have been found for female (SPL increase of 11.1 dB) and male (9.3 dB increase) speakers for every doubling of P_{sg}', with an average r^2 of .69 for the relationship (Björklund & Sundberg, 2016). This strong logarithmic relationship provided the necessary participant-specific mapping from P_{sg}' to SPL.

Thus, the third performance metric was the PI_{95} for P_{sg}'-derived SPL that could be directly compared with the uncertainty reported by Švec et al. (2005). Given that the PI_{95} for ACC-derived P_{sg}' was already computed, we only needed to multiply this interval by an appropriate conversion factor. However, because the relationship between P_{sg}' (in cm H_2O) and SPL (in dB SPL) is logarithmic, the relationship between the two was first linearized by converting P_{sg}' to a decibel scale using $20\log_{10}(P_{sg}')$ to yield units of dB re cm H_2O. Figure 2 illustrates this linearization for the strong SPL–P_{sg}' relationship (r^2 = .87) for female participant F1 in the current study (average r^2 = .85 across participants). The linear regression slope of 1.76 dB/dB was thus used as the multiplier to convert the PI_{95} for ACC-derived P_{sg}' to the PI_{95} for P_{sg}'-derived SPL. For this participant, SPL increased by 10.5 dB for every doubling (+6 dB) of P_{sg}', which compares well with the reported values in the literature.

Results

Correlation Between ACC Amplitude and P_{sg}'

Figure 3 displays scatter plots illustrating the relationships for female participant F3 between P_{sg}' and ACC amplitude and between MIC amplitude and ACC amplitude, as well as the corresponding PI_{95} for estimating these measures with the ACC. Color-coding and symbols display the effects of vowel type and fundamental frequency on these relationships. In particular, the relationship between P_{sg}' and ACC amplitude remains relatively robust to vowel and pitch differences compared with the relationship between MIC amplitude and ACC amplitude.

Figures 3d and 3f show that linear trends are exhibited within specific vowel types and pitch levels but that, when pooled together, the variation (PI_{95}) changes from ±1.85 cm H_2O for the P_{sg}'–ACC relationship to ±6.13 dB for the MIC–ACC relationship. The bifurcation in Figure 3d is largely due to inherent vowel differences in supraglottal filter or radiation characteristics; in particular, vowel /a/

Figure 2. Illustration of the relationship between acoustic sound pressure level (SPL) and estimated subglottal pressure (P_{sg}') for female participant F1: (a) Logarithmic relationship between SPL and P_{sg}' in cm H_2O. (b) Linearized relationship between SPL and P_{sg}' converted to a decibel scale.

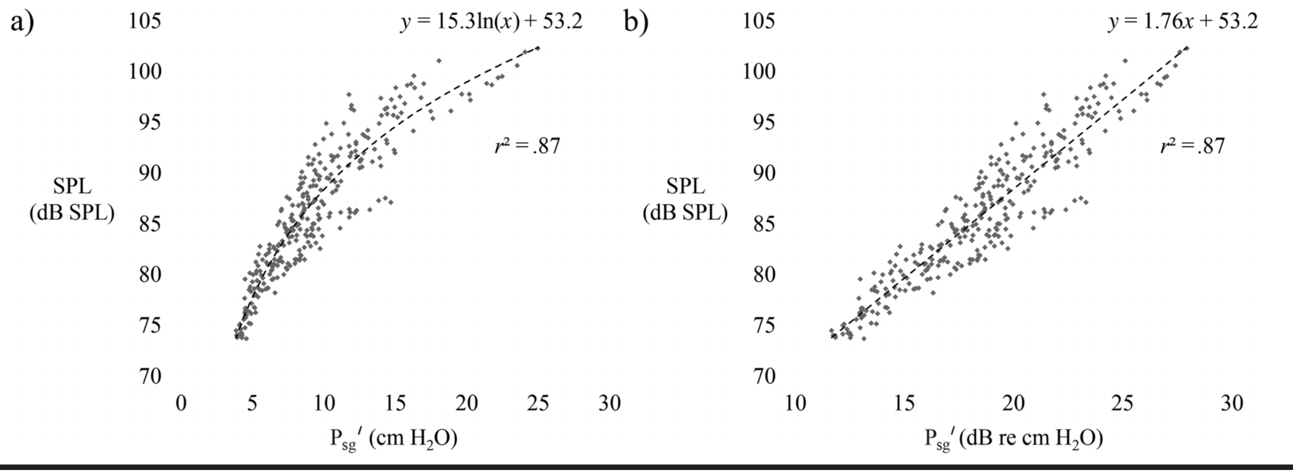
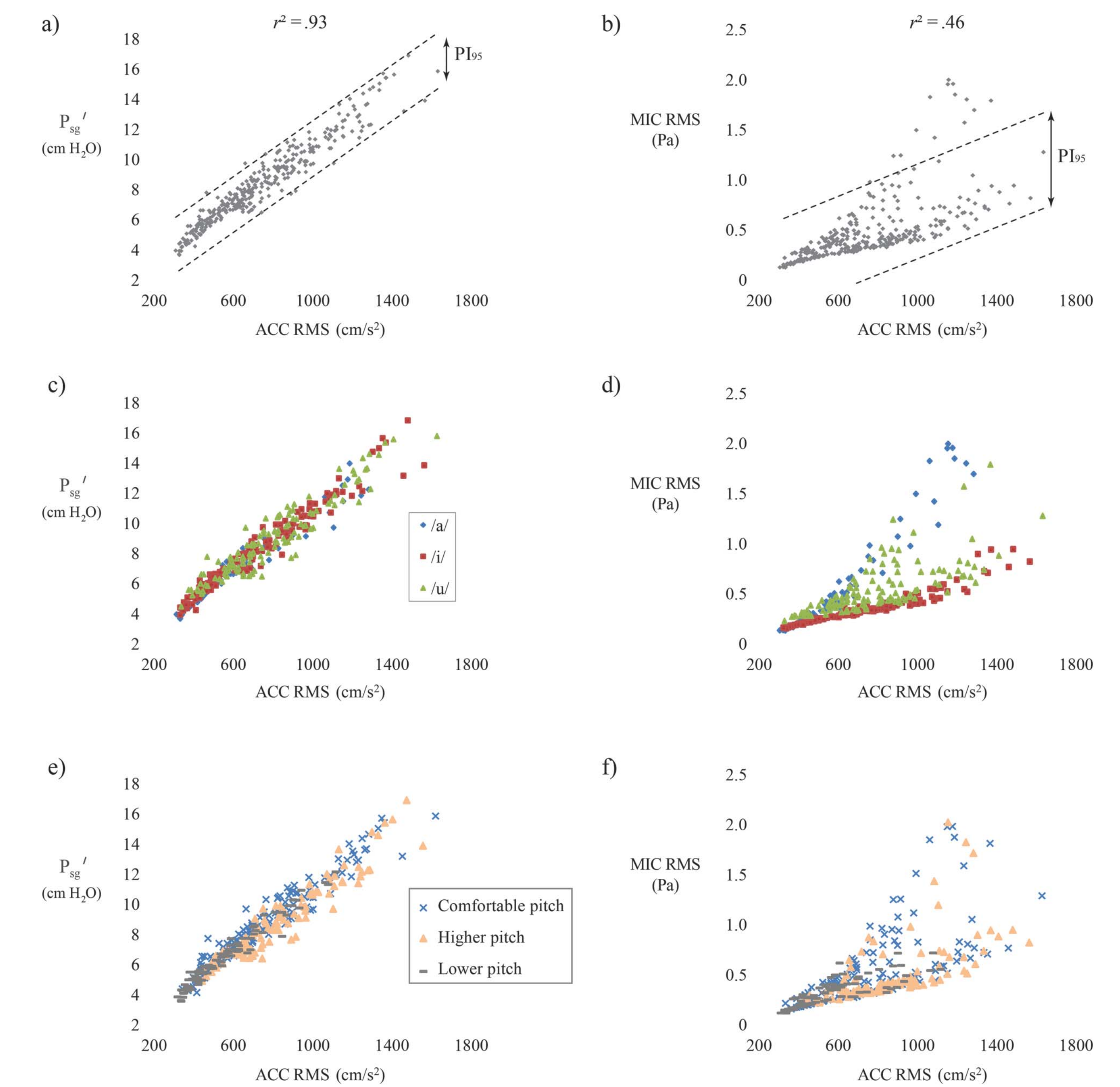


Figure 3. Scatter plots illustrating the estimated subglottal pressure (P_{sg}')–accelerometer (ACC) and microphone (MIC)–ACC relationships for female participant F3. Shown are the overall relationships (including 95% prediction interval [PI_{95}]) between (a) P_{sg}' and ACC root-mean-square (RMS) amplitude and (b) MIC- and ACC-signal RMS amplitudes. Subsequent plots are coded to show effects of (c, d) vowel context and (e, f) pitch condition.



produced at a comfortable pitch level exhibits a steeper slope between MIC and ACC amplitudes than the corresponding slope exhibited by vowel /i/. Due to the subglottal placement of the ACC sensor, vowel dependencies are less prominent in the ACC amplitude and thus lead to a closer correspondence with P_{sg}' (see Figure 3c). The higher variance in the MIC–ACC relationship for the vowel /u/ may be due to challenges in producing consistent lip openings for /u/ productions of varying loudness and pitch between /p/ consonants, which further argues for the use of ACC amplitude as a more stable measure of vocal intensity relative to MIC-derived SPL.

Figure 4 displays r^2 values for the participant-specific P_{sg}'–ACC and MIC–ACC relationships. Overall, P_{sg}' and ACC-signal amplitude exhibited a high coefficient of determination (r^2 = .68–.93) for each speaker, pooling across all

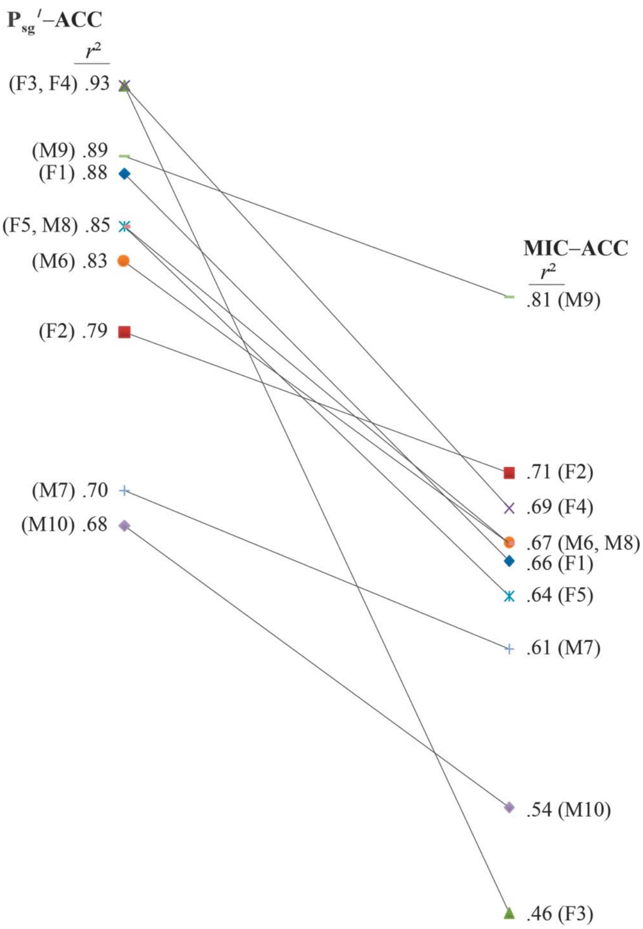


Figure 4. Overall r^2 values between estimated subglottal pressure (P_{sg}') and accelerometer (ACC)-signal amplitude, and between microphone (MIC) sound pressure level and ACC-signal amplitude for each participant, pooling across vowel and pitch conditions.

vowel and pitch conditions. These r^2 values were consistently higher than those obtained for the MIC–ACC relationship (r^2 = .46–.81) for each participant. Of note, r^2 values did not change significantly when all pressure plateaus—including those with peak-to-peak amplitudes above 0.5 cm H_2O—were included in the analysis.

Effect of Vowel Context and Pitch Condition

Table 1 reports the effects of vowel context and pitch condition in terms of r^2 values for the P_{sg}'–ACC and MIC–ACC relationships. Correlations were available for nine pitch–vowel combinations because three vowels and three pitch levels were elicited from each participant. To summarize the effect of vowel context, data from all three vowels were pooled within each pitch condition to yield three r^2 values. These three r^2 values were averaged to yield a summary statistic reflecting the robustness of a relationship to changes in vowel context.

Table 1. The effect of vowel context and pitch condition on the reported coefficient of determination (r^2) for the relationships between estimated subglottal pressure (P_{sg}') and accelerometer (ACC) and microphone and accelerometer.

Participant ID	Pooling vowel contexts		Pooling pitch conditions	
	P_{sg}'–ACC	MIC–ACC	P_{sg}'–ACC	MIC–ACC
F1	.88	.63	.84	.57
F2	.92	.83	.80	.77
F3	.93	.52	.93	.78
F4	.94	.69	.93	.86
F5	.87	.60	.92	.92
M6	.81	.67	.85	.81
M7	.79	.75	.76	.67
M8	.90	.70	.86	.81
M9	.90	.87	.88	.84
M10	.64	.49	.71	.72

Note. Pooling vowel contexts: Data within vowel contexts are pooled within each pitch condition, and r^2 is then averaged across the three pitch conditions. Pooling pitch conditions: Data within pitch condition are pooled within each vowel context, and r^2 is then averaged across the three vowel contexts.

With the exception of participant F2, r^2 remained similar for the P_{sg}'–ACC relationship when data were pooled separately across vowel contexts and across pitch conditions. This result indicated that neither vowel nor fundamental frequency had a strong impact on the variation in the P_{sg}'–ACC relationship. In contrast, there were notable differences between r^2 in all but one participant (M9) when we examined MIC–ACC relationships in terms of vowel and pitch contexts. Changes in pitch condition led to more uncertainty in six participants, whereas vowel context was linked to more uncertainty in three participants.

Comparison of Uncertainty in P_{sg}'–ACC and MIC–ACC Relationships

Table 2 shows the PI_{95} for estimating P_{sg}' within each participant using ACC-signal amplitude. On average, the PI_{95} (Column 2) was ±2.53 cm H_2O across the 10 participants, ranging from ±1.70 to ±3.74 cm H_2O. In order to compare these values with the known uncertainty in predicting SPL using ACC-signal amplitude, we transformed P_{sg}' estimates to a logarithmic scale in Column 3. Those values were then multiplied by respective slopes computed from the regression of SPL onto logarithmic P_{sg}' (recall Figure 2) to yield PI_{95} values in Column 4. Overall, there was consistently less uncertainty associated with P_{sg}'-derived SPL than the uncertainty associated with ACC-derived SPL (Column 5).

Figure 5 represents an attempt to find a single overall mapping between P_{sg}' and ACC-signal amplitude and between MIC and ACC levels across all participants. The coefficient of determination found for the P_{sg}'–ACC regression equation (r^2 = .63) was weaker than the lowest

Table 2. Average 95% prediction interval for predicting subglottal pressure (P_{sg}') and acoustic sound pressure level (SPL) using accelerometer (ACC)-signal amplitude.

Participant ID	ACC-derived P_{sg}' (cm H_2O)	ACC-derived P_{sg}' (dB)	P_{sg}'-derived SPL (dB)	ACC-derived SPL (dB)
F1	±3.05	±2.40	±4.22	±7.01
F2	±2.12	±2.30	±3.84	±4.34
F3	±1.85	±1.37	±2.23	±6.13
F4	±2.43	±3.36	±6.34	±8.67
F5	±2.43	±2.25	±4.76	±6.38
M6	±2.03	±2.78	±6.67	±6.86
M7	±2.72	±3.23	±4.60	±4.88
M8	±3.22	±2.82	±4.49	±5.20
M9	±1.70	±2.35	±3.66	±4.50
M10	±3.74	±5.72	±7.10	±7.44
Average	±2.53	±2.86	±4.79	±6.14

coefficient obtained for a single participant, pointing toward the potential need for requiring participant-specific calibrations to use a regression equation to predict P_{sg}' from ACC-signal amplitude. The overall relationship between MIC and ACC signal levels yielded an expected low r^2 of .30.

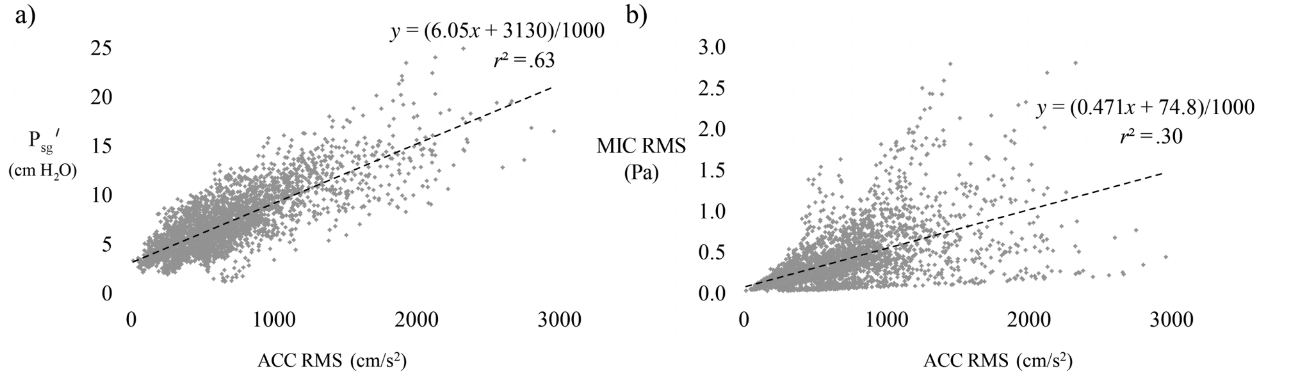
Discussion

This study examined the relationship between a subglottal neck-surface ACC signal and P_{sg}' during sustained vowel production. It was hypothesized that ACC-signal amplitude would correlate well with P_{sg}' due to its role in capturing neck-surface vibrations below the glottis, which are generated largely by radiated subglottal aeroacoustic energy. Previous work has also supported this view, with reliable estimates of other aerodynamic measurements (e.g., AC flow and maximum flow declination rate) successfully obtained through inverse filtering of the ACC waveform (Zañartu, Ho, et al., 2013). However, that inverse filtering technique does not explicitly model or provide for the estimation of subglottal air pressure from the ACC signal. Thus, the current work begins to address this missing link by obtaining empirical relationships between ACC-signal amplitude and average P_{sg}' to ultimately aid in extending numerical models of voice production. Furthermore, it attempted to explore the derivation of a voice-production measure that could be robust to variations that would be exhibited during continuous speech.

An empirical relationship has been previously obtained between ACC-signal amplitude and SPL (dB/dB scale), but with an inherent average variability of ±6 and ± 5 dB, respectively, for male and female adult speakers (Švec et al., 2005). We hypothesized that much of this uncertainty was due to rapidly changing vocal-tract resonances during running speech. In other words, an individual can maintain one intensity level at the source, but the level at the lip opening will vary depending on vocal-tract shape. Furthermore, we hypothesized that the estimation of subglottal air pressure (instead of SPL) from ACC-signal amplitude would exhibit less uncertainty than would estimation of SPL from ACC-signal amplitude.

As hypothesized, coefficients of determination between ACC-signal amplitude and P_{sg}' were strong, ranging from r^2 = .68 to .93. We attempted to find a robust single relationship across all participants; however, this resulted in a weaker coefficient (r^2 = .63), likely due to speaker-specific differences in vibratory output on the basis of anatomical

Figure 5. Relationships between (a) estimated subglottal pressure (P_{sg}') and accelerometer (ACC) root-mean-square (RMS) amplitude across all participants and (b) microphone (MIC) RMS amplitude and ACC RMS amplitude.



variations of skin stiffness, tracheal length, and laryngeal anatomy (Zañartu, Ho, et al., 2013). To translate these results to an ambulatory ACC context, individualized calibration sessions may be necessary to robustly estimate subglottal pressure. Future work calls for a better theoretical understanding of the different participant-specific slopes and offsets of the P_{sg}'–ACC regression lines.

In addition, both vowel and fundamental-frequency contexts appeared to contribute to the uncertainty of estimating SPL using ACC-signal amplitude. This was evident when we examined the data categorized by vowel type or fundamental frequency. The same is not true when analogous data are plotted for P_{sg}' and ACC amplitude. In these instances, the data are positioned closely together. Further, the average of the correlations across vowels and fundamental frequencies were largely unchanged in the P_{sg}'–ACC relationship in eight of the 10 participants, supporting the notion that subglottal resonances remain stable across speech contexts. The variability in average correlations for two participants (F2 and M10) may be due to the fact that these participants had the lowest overall correlations to begin with. In addition, SPL results are tempered by the use of a directional microphone that is known to exhibit proximity effects at close distances (boosting low-frequency energy) and a nonflat frequency response (gradually attenuating energy below 500 Hz), although these properties counteract each other in practice.

These results support the conclusion that the P_{sg}'–ACC relationship contains a lesser degree of variability than the acoustic MIC–ACC relationship in typical voices. The uncertainty associated with ACC-derived P_{sg}' was consistently lower than that associated with ACC-derived SPL. This suggests that estimating P_{sg}' in a continuous speech context may yield a vocal-function parameter that is more robust than the current estimation of SPL in ambulatory contexts. However, future work is needed to investigate the translation of ACC-derived P_{sg}' measured using /p/–vowel syllable strings to continuous speech. Using invasive techniques (tracheal puncture, esophageal balloon, etc.) may be necessary in a small number of participants to directly measure subglottal pressure simultaneously with neck-surface acceleration to gain additional insight into differences between vowel and continuous speech contexts.

For this study, a new strategy for removing undesirable "peaky" pressure signals was established in hopes of standardizing this process. Filtering the data in a consistent way combined the traditional approach of visually accepting flat signals with a quantitative method of removing undesirable pressure plateaus. However, when all pressure plateaus were included for analysis, correlations did not change significantly. This result is likely due to the specific instructions and training given to participants that optimized their ability to achieve acceptable plateaus. It is posited that including only the best-performed P_{sg}' estimations supports the assertion that the relationship between the ACC and P_{sg}' is indeed a strong one.

This study provides preliminary support for estimating P_{sg}' using an ACC sensor in vocally healthy individuals. More work is needed, though, to determine whether the same paradigm can be applied to speakers phonating in nonmodal voice qualities (e.g., roughness, breathiness, strain) and to individuals who have been diagnosed with vocal pathologies affecting the voice-production mechanism. It is possible that the relationship between P_{sg}' and ACC-signal amplitude obtained in this study will not remain as strong for certain voice types. For instance, in a voice characterized by a strained vocal quality, the buildup of P_{sg}' to initiate voice may not be commensurate with the vibratory output of the larynx. The ACC signal, in turn, may underestimate the amount of P_{sg}' present. Thus, future work for determining the effectiveness of this methodology in individuals with atypical vocal characteristics is warranted.

Conclusion

Findings suggest that estimates of subglottal pressure during phonation may be extracted using a simple signal level of a neck-mounted accelerometer sensor. Estimating subglottal pressure was more accurate than estimating acoustic sound pressure level using subglottal neck-surface acceleration. Acquiring these estimates would be valuable, because aerodynamic measures such as P_{sg}' have been associated with disorders related to voice use (e.g., vocal hyperfunction; Hillman et al., 1989). Moving forward, the ability to estimate subglottal pressure noninvasively and in a persistent, long-term manner may afford clinicians the opportunity to obtain a more comprehensive picture of a patient's voice use to better diagnose and treat aberrant vocal behaviors. Future work is warranted to explore the salience of the proposed technique in individuals with voice disorders.

Acknowledgments

This work was supported by Voice Health Institute and National Institute on Deafness and Other Communication Disorders Grant R33 DC011588 (awarded to Robert E. Hillman). Contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health. The authors thank J. T. Heaton for scientific discussions and feedback on earlier versions of the manuscript. Portions of this work were presented in the Master's thesis of the first author (Fryd, 2015).

References

Björklund, S., & Sundberg, J. (2016). Relationship between subglottal pressure and sound pressure level in untrained voices. *Journal of Voice, 30,* 15–20.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.3.80) [Computer program]. Retrieved from http://www.praat.org/

Bottalico, P., & Astolfi, A. (2012). Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms. *The Journal of the Acoustical Society of America, 131,* 2817–2827.

Cheyne, H. A. (2002). *Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge.

Cheyne, H. A., Hanson, H. M., Genereux, R. P., Stevens, K. N., & Hillman, R. E. (2003). Development and testing of a portable vocal accumulator. *Journal of Speech, Language, and Hearing Research, 46,* 1457–1467.

Colton, R. H., Casper, J. K., & Leonard, R. (2006). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (3rd ed.). Baltimore, MD: Lippincott, Williams & Wilkins.

Cranen, B., & Boves, L. (1985). Pressure measurements during speech production using semiconductor miniature pressure transducers: Impact on models for speech production. *The Journal of the Acoustical Society of America, 77,* 1543–1551.

Döllinger, M., Kobler, J., Berry, D. A., Mehta, D. D., Luegmair, G., & Bohr, C. (2011). Experiments on analysing voice production: Excised (human, animal) and *in vivo* (animal) approaches. *Current Bioinformatics, 6,* 286–304.

Fryd, A. S. (2015). *Estimating subglottal pressure from neck-surface acceleration* (Unpublished master's dissertation). MGH Institute of Health Professions, Boston, MA.

Hartl, D. M., Hans, S., Vaissière, J., Riquet, M., & Brasnu, D. F. (2001). Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *Journal of Voice, 15,* 351–361.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97,* 3099–3111.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32,* 373–392.

Hillman, R. E., Montgomery, W. W., & Zeitels, S. M. (1997). Appropriate use of objective measures of vocal function in the multidisciplinary management of voice disorders. *Current Opinion in Otolaryngology & Head and Neck Surgery, 5,* 172–175.

Hixon, T. J. (1972). Some new techniques for measuring the biomechanical events of speech production: One laboratory's experiences. *ASHA Reports, 7,* 68–103.

Holmberg, E. B., Doyle, P., Perkell, J. S., Hammarberg, B., & Hillman, R. E. (2003). Aerodynamic and acoustic voice measurements of patients with vocal nodules: Variation in baseline and changes across voice therapy. *Journal of Voice, 17,* 269–282.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America, 84,* 511–529.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1989). Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch. *Journal of Voice, 3,* 294–305.

Jiang, J., & Stern, J. (2004). Receiver operating characteristic analysis of aerodynamic parameters obtained by airflow interruption: A preliminary report. *Annals of Otology, Rhinology & Laryngology, 113,* 961–966.

Lindstrom, F., Waye, K. P., Södersten, M., McAllister, A., & Ternström, S. (2011). Observations of the relationship between noise exposure and preschool teacher voice usage in day-care center environments. *Journal of Voice, 25,* 166–172.

Löfqvist, A., Carlborg, B., & Kitzing, P. (1982). Initial validation of an indirect measure of subglottal pressure during vowels. *The Journal of the Acoustical Society of America, 72,* 633–635.

Mehta, D. D., Van Stan, J. H., & Hillman, R. E. (2016). Relationships between vocal function measures derived from an acoustic microphone and a subglottal neck-surface accelerometer. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24,* 659–668.

Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., . . . Hillman, R. E. (2015). Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in Bioengineering and Biotechnology, 3,* 155.

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

Nacci, A., Fattori, B., Mancini, V., Panicucci, E., Ursino, F., Cartaino, F. M., & Berrettini, S. (2013). The use and role of the Ambulatory Phonation Monitor (APM) in voice assessment. *Acta Otorhinolaryngologica Italica, 33,* 49–55.

Netsell, R., Lotz, W., & Shaughnessy, A. L. (1984). Laryngeal aerodynamics associated with selected voice disorders. *American Journal of Otolaryngology, 5,* 397–403.

Plant, R. L., & Hillel, A. D. (1998). Direct measurement of subglottic pressure and laryngeal resistance in normal subjects and in spasmodic dysphonia. *Journal of Voice, 12,* 300–314.

Popolo, P. S. (2007). *Relating vocal fold vibration amplitude to skin acceleration level on the anterior neck* Unpublished doctoral dissertation, The University of Iowa, Iowa City.

Ramig, L. O., & Dromey, C. (1996). Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with Parkinson disease. *Journal of Speech and Hearing Research, 39,* 798–807.

Rosenthal, A. L., Lowell, S. Y., & Colton, R. H. (2014). Aerodynamic and acoustic features of vocal effort. *Journal of Voice, 28,* 144–153.

Rothenberg, M. (1973). A new inverse-filtering technique for deriving glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America, 53,* 1632–1645.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology, 22,* 212–226.

Speyer, R. (2008). Effects of voice therapy: A systematic review. *Journal of Voice, 22,* 565–580.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Sundberg, J., Scherer, R., Hess, M., Müller, F., & Granqvist, S. (2013). Subglottal pressure oscillations accompanying phonation. *Journal of Voice, 27,* 411–421.

Švec, J. G., Titze, I. R., & Popolo, P. S. (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *The Journal of the Acoustical Society of America, 117,* 1386–1394.

Tanaka, S., & Gould, W. J. (1983). Relationships between vocal intensity and noninvasively obtained aerodynamic parameters in normal subjects. *The Journal of the Acoustical Society of America, 73,* 1316–1321.

Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America, 85,* 901–906.

Titze, I. R. (1992). Vocal efficiency. *Journal of Voice, 6,* 135–138.

Titze, I. (2013). Quantifying vocal efficiency and economy—How can computation augment clinical assessment? *Proceedings of Meetings on Acoustics, 19,* 060244.
Titze, I. R., & Hunter, E. J. (2011). Feasibility of measurement of a voice range profile with a semi-occluded vocal tract. *Logopedics Phoniatrics Vocology, 36,* 32–39.
Titze, I. R., & Hunter, E. J. (2015). Comparison of vocal vibration-dose measures for potential-damage risk criteria. *Journal of Speech, Language, and Hearing Research, 58,* 1425–1439.
Titze, I. R., Maxfield, L., & Palaparthi, A. (2016). An oral pressure conversion ratio as a predictor of vocal efficiency. *Journal of Voice, 30,* 398–406.
Titze, I. R., Švec, J. G., & Popolo, P. S. (2003). Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *Journal of Speech, Language, and Hearing Research, 46,* 919–932.
van den Berg, J. (1956). Direct and indirect determination of the mean subglottic pressure: Sound level, mean subglottic pressure, mean air flow, "subglottic power" and "efficiency" of a male voice for the vowel (a). *Folia Phoniatrica, 8,* 1–24.
Van Stan, J. H., Mehta, D. D., Zeitels, S. M., Burns, J. A., Barbu, A. M., & Hillman, R. E. (2015). Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects. *Annals of Otology,* Rhinology & Laryngology, *124,* 864–874.
Wokurek, W., & Madsack, A. (2011). Acceleration sensor based estimates of subglottal resonances: Short vs. long vowels. In P. Cosi, R. De Mori, G. Di Fabbrizio, & R. Pieraccini (Eds.), *Interspeech 2011, 12th Annual Conference of the International Speech Communication Association* (pp. 2821–2824). Baixas, France: International Speech Communication Association.
Wokurek, W., & Pützer, M. (2009). Accelleration sensor measurements of subglottal sound pressure for modal and breathy phonation quality. In C. Manfredi (Ed.), *6th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications* (pp. 153–156). Firenze, Italy: Firenze University Press.
Zañartu, M., Espinoza, V., Mehta, D. D., Van Stan, J. H., Cheyne, H. A., II, Ghassemi, M., . . . Hillman, R. E. (2013). Toward an objective aerodynamic assessment of vocal hyperfunction using a voice health monitor. In C. Manfredi (Ed.), *8th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications* (pp. 167–170). Firenze, Italy: Firenze University Press.
Zañartu, M., Galindo, G. E., Erath, B. D., Peterson, S. D., Wodicka, G. R., & Hillman, R. E. (2014). Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *The Journal of the Acoustical Society of America, 136,* 3262–3271.
Zañartu, M., Ho, J. C., Mehta, D. D., Hillman, R. E., & Wodicka, G. R. (2013). Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Transactions on Audio, Speech, and Language Processing, 21,* 1929–1939.
Zeitels, S. M., Blitzer, A., Hillman, R. E., & Anderson, R. R. (2007). Foresight in laryngology and laryngeal surgery: A 2020 vision. *Annals of Otology, Rhinology & Laryngology, 116*(Suppl. 198), 1–16.
Zeitels, S. M., Hillman, R. E., Franco, R. A., & Bunting, G. W. (2002). Voice and treatment outcome from phonosurgical management of early glottic cancer. *Annals of Otology, Rhinology & Laryngology, 111*(Suppl. 190), 1–20.
Zeitels, S. M., Hochman, I., & Hillman, R. E. (1998). Adduction arytenopexy: A new procedure for paralytic dysphonia with implications for implant medialization. *Annals of Otology, Rhinology & Laryngology, 107*(Suppl. 173), 1–24.
Zeitels, S. M., Lopez-Guerra, G., Burns, J. A., Lutch, M., Friedman, A. M., & Hillman, R. E. (2009). Microlaryngoscopic and office-based injection of bevacizumab (Avastin) to enhance 532-nm pulsed KTP laser treatment of glottal papillomatosis. *Annals of Otology, Rhinology & Laryngology, 118*(Suppl. 201), 1–24.