



# HHS Public Access

Author manuscript

*IEEE/ACM Trans Audio Speech Lang Process.* Author manuscript; available in PMC 2016 April 08.

Published in final edited form as:

*IEEE/ACM Trans Audio Speech Lang Process.* 2016 April ; 24(4): 659–668. doi:10.1109/TASLP.

2016.2516647

## Relationships between vocal function measures derived from an acoustic microphone and a subglottal neck-surface accelerometer

**Daryush D. Mehta [Member, IEEE],**

Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston MA 02114 USA, Department of Surgery, Harvard Medical School, Boston, MA 02115 USA, and the Institute of Health Professions, Massachusetts General Hospital, Boston, Massachusetts 02129 USA (mehta.daryush@mgh.harvard.edu)

**Jarrad H. Van Stan, and**

Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston MA 02114 USA and the Institute of Health Professions, Massachusetts General Hospital, Boston, Massachusetts 02129 USA (jvanstan@mghihp.edu)

**Robert E. Hillman**

Center for Laryngeal Surgery & Voice Rehabilitation and Institute of Health Professions, Massachusetts General Hospital, Boston MA 02114 USA and Surgery and Health Sciences & Technology, Harvard Medical School, Boston, MA 02115 (hillman.robert@mgh.harvard.edu)

### Abstract

Monitoring subglottal neck-surface acceleration has received renewed attention due to the ability of low-profile accelerometers to confidentially and noninvasively track properties related to normal and disordered voice characteristics and behavior. This study investigated the ability of subglottal neck-surface acceleration to yield vocal function measures traditionally derived from the acoustic voice signal and help guide the development of clinically functional accelerometer-based measures from a physiological perspective. Results are reported for 82 adult speakers with voice disorders and 52 adult speakers with normal voices who produced the sustained vowels /a/, /i/, and /u/ at a comfortable pitch and loudness during the simultaneous recording of radiated acoustic pressure and subglottal neck-surface acceleration. As expected, timing-related measures of jitter exhibited the strongest correlation between acoustic and neck-surface acceleration waveforms ( $r = 0.99$ ), whereas amplitude-based measures of shimmer correlated less strongly ( $r = 0.74$ ). Additionally, weaker correlations were exhibited by spectral measures of harmonics-to-noise ratio ( $r = 0.69$ ) and tilt ( $r = 0.57$ ), whereas the cepstral peak prominence correlated more strongly ( $r = 0.90$ ). These empirical relationships provide evidence to support the use of accelerometers as effective complements to acoustic recordings in the assessment and monitoring of vocal function in the laboratory, clinic, and during an individual's daily activities.

### Index Terms

vocal function analysis; neck-surface accelerometer; vocal perturbation; ambulatory voice monitoring; cepstral peak prominence; harmonics-to-noise ratio

## I. Introduction

BODY surface vibrations generated during speaking often provide robust signals that can be related to the underlying physiological mechanisms of voice and speech production. Accelerometer (ACC) sensors can measure these signals by taking advantage of the piezoelectric effect to transduce mechanical forces into electrical signals. Since the 1960s, investigators have employed ACCs to supplement or replace acoustic microphone (MIC) recordings for selected applications in order to obtain estimates of parameters that are minimally affected by unwanted acoustic interference (e.g. environmental noise, speech of others, etc.). When placed appropriately, in addition to sensing phonation, ACCs can provide spectral features related to chest wall vibration [1], nasal resonances [2], [3], and subglottal resonances [4].

Recently, multiple types of nonacoustic body sensors have been shown to complement each other in speech communication systems that require robust performance in the presence of high noise environments [5]. ACCs placed on the anterior neck below the larynx are particularly well suited for measuring phonation because of their relative insensitivity to the time-varying speech signal and background noise levels [6], thus providing potentially significant benefits in the study of normal and disordered vocal function. In fact, when compared to MICs, ACCs are more immune to environmental noise artifacts [6]. Furthermore, when positioned on the anterior neck surface during voice production, ACCs may measure components related to tissue-to-tissue transmission of vocal fold collision forces through the thyroid cartilage and air-to-tissue transmission of aerodynamic energy through the trachea [7], [8]. The relative contribution of these two components may play a critical role in how ACC-derived measures complement their MIC-derived counterparts in the characterization of normal and disordered phonation.

In terms of voice source characterization, anterior neck-surface acceleration at the tracheal level has been studied in speakers with both normal and disordered voices to derive features related to average fundamental frequency ( $f_0$ ) [9], [10], instantaneous  $f_0$  [11], sound pressure level [12], voice activity detection [13], and glottal airflow features [14]. Robust estimation of vocal  $f_0$  has been the primary motivation for employing neck-placed ACCs, particularly in noisy environments [9], [10] and in breathy speech contexts when electroglottography fails to register a signal due to reduced vocal fold contact [15]. The Spearman correlation between MIC- and ACC-based  $f_0$  has been reported to range from 0.73 to 0.92 during continuous speech in four speakers with normal voices [16]. Anecdotal evidence of short-term variation in the  $f_0$  has demonstrated that jitter values as measured by an MIC are similar to jitter derived from an ACC signal [17].

The field of ambulatory voice monitoring or *voice dosimetry* has heavily relied upon the estimation of  $f_0$  and sound pressure level from a neck-mounted ACC with the primary objective of quantifying the accumulated impact of prolonged voice use by speakers in occupations with high vocal demands [18], [19], [13], [20], [21]. ACC-based recordings are well suited for ambulatory monitoring as the ACC sensor is robust in the context of background noise and preserves speaker confidentiality when placed below the larynx (i.e., the sensor does not capture intelligible speech). However, there are limitations in the use of

accelerometry to estimate sound pressure level as the short-time energy in the ACC signal appears to only correlate with the acoustic sound pressure level to a particular degree of uncertainty that approaches  $\pm 6$  dB [12]. Nevertheless, it is hoped that ACC-based voice monitoring systems will provide complementary information with data obtained from in-laboratory MIC recordings, especially since certain voice disorders are associated with aberrant patterns of daily voice use [22].

Acoustic measures based on cycle-to-cycle perturbation (e.g., jitter and shimmer) and signal-to-noise (e.g., harmonics-to-noise, cepstral peak prominence, etc.) levels are often used clinically to objectively assess the impact of voice disorders on vocal function. Acoustic measures of perturbation such as jitter have historical roots as indicators of “the physical processes of speech production” (p. 344) [23], and spectral measures of noise and tilt have been used to characterize glottal closure patterns [24]. In the cepstral domain, variations in acoustic measures have been obtained to study trading/compensatory relationships between vocal fold vibratory characteristics such as asymmetry, speed quotient, and incomplete glottal closure [25]. Based on the advantages described (e.g., relative immunity to environmental noise), use of the high-bandwidth ACC signal has the potential to enhance these types of voice assessment. Aside from some case studies, however, which have reported that the MIC signal tends to exhibit approximately twice as much shimmer as in simultaneously recorded ACC signals [17], little attention has been focused on whether the neck-surface ACC signal can be used to estimate analogous parameters.

This study was motivated by the desire to extract more information from the neck-surface ACC signal, especially characteristics that may be translated from vocal function measures used in clinical voice assessment. The specific purpose was to determine the extent to which vocal function measures extracted from the subglottal neck-surface ACC signal are related to analogous measures derived from the MIC signal in speakers with and without voice disorders. Sustained vowel production was thus analyzed for three categories of vocal function measures: (1) time-domain perturbation (jitter, shimmer, harmonics-to-noise ratio), (2) spectral characteristics (harmonics-to-noise ratio, spectral tilt), and (3) cepstral properties (cepstral peak prominence). It is acknowledged that using MIC-derived measures as reference metrics may be considered imperfect as researchers continue to elucidate the functional significance of various objective measures for clinical voice assessment [26]. However, the ongoing development of clinically significant acoustic measures, particularly those based on the cepstrum, shows promise for distinguishing voice qualities and classifying patients from vocally-normal speakers, e.g., [27].

It is hypothesized that timing-related measures will compare well between the ACC and MIC domains due to high correlations for average  $f_0$  in the literature and the theoretical basis that similar information from phonatory cycles radiates through the neck tissue and through the vocal tract and out of the mouth. Amplitude-based measures are expected to exhibit a decreased correlation between ACC and MIC signals due to larger variances across subjects of the ACC waveshapes. The lowpass-filter quality of the neck frequency response has been observed to be  $-8.4$  dB per octave (dB/oct) for individuals with normal neck tissue [28] and  $-8.8$  dB/oct across both laryngectomee patients and normal subjects [29]. Thus the spectral tilt of the ACC signal is hypothesized to have a statistically significant bias on the

order of 8 dB/oct when compared to the spectral tilt of the MIC signal. Since the cepstral peak prominence (CPP) is an integrative measure of perturbation, harmonics-to-noise ratio [30], [31], [32], and waveshape differences, it is expected that CPP will moderately correlate between ACC and MIC domains. Glottal turbulence noise, if present in the MIC signal, is hypothesized to be significantly attenuated in the neck-surface ACC signal, thus decreasing the variance of CPP measures in the ACC domain.

## II. Methods

### A. Subject enrollment

The study sample consisted of 134 adult speakers: 52 subjects (47 female, 5 male) with normal voices and 82 subjects (69 female, 13 male) diagnosed with a voice disorder such as muscle tension dysphonia or having benign vocal fold lesions such as nodules and/or polyps. The average age (mean  $\pm$  SD) of subjects with normal voices was  $27.3 \pm 11.4$  years for the female group and  $29.4 \pm 8.4$  years for the male group. The average age of subjects with voice disorders was  $32.4 \pm 14.5$  years for the female group and  $38.4 \pm 11.9$  years for the male group. In the group with voice disorders, 46 of the subjects were assessed during multiple visits throughout the course of treatment (e.g., before and after laryngeal surgery or voice therapy): 37 subjects were assessed twice, 7 subjects were assessed three times, and 2 subjects were assessed four times. Thus, data were acquired over 191 total sessions.

### B. Data collection

Figure 1 shows the data acquisition setup. Subjects were enrolled in a larger study on smartphone-based ambulatory voice monitoring whose in-laboratory protocol called for the simultaneous acquisition of acoustics, electroglottography, subglottal neck-surface acceleration, and aerodynamic estimates of oral airflow and subglottal pressure (pneumotachograph mask system) [33]. The current study focused on obtaining vocal function metrics from the MIC and ACC data. Each subject was instructed to sustain three vowels (/a/, /i/, /u/), each for 2–5 s at a comfortable pitch and loudness.

The MIC signal was recorded using a head-mounted condenser MIC with a cardioid pattern (Model MKE104, Sennheiser electronic GmbH, Wedemark-Wennebostel, Germany). The MIC was situated approximately 4 cm from the lips at a 45-degree azimuth. The MIC signal was input to a preamplifier (Model 302 Dual Microphone Preamplifier, Symetrix, Inc., Mountlake Terrace, WA), followed by preconditioning electronics (CyberAmp Model 380, Axon Instruments, Inc., Union City, CA) for gain control and anti-alias filtering at a 3 dB cutoff frequency of 8 kHz. The analog signal was digitized at a 20 kHz sampling rate, 16-bit quantization, and  $\pm 10$  V dynamic range (Digidata Model 1440A, Axon Instruments, Inc.).

The ACC consisted of a miniature piezo-ceramic vibration transducer (BU-27135, Knowles Electronics) with unidirectional sensitivity in the axial dimension and dimensions  $7.92 \text{ mm} \times 5.59 \text{ mm} \times 4.14 \text{ mm}$ . The ACC has a linear frequency response from 20 Hz to 20 kHz and was wired to a three-conductor cable and mounted on a flexible silicone pad with a durable silicone sealant and epoxy. This sensor mounting was necessary to provide for a durable assembly that was also used for ambulatory monitoring of subjects in context of the larger

study [20]. The ACC assembly (sensor mounted on the silicone pad) was calibrated to physical units of acceleration ( $\text{cm/s}^2$ ) by sending a known stimulus to a mechanical shaker. The sensor was affixed to the subject's anterior neck-skin surface halfway between the thyroid prominence and the suprasternal notch along the midsagittal axis using hypoallergenic double-sided tape (Model 2181, 3M, Maplewood, MN).

The ACC signal was recorded with a sampling rate of 11.025 kHz (16-bit quantization) on a Google/Samsung Nexus S smartphone that allowed for programmable gain control prior to input into a sigma-delta modulation audio codec (WM8994; Wolfson Microelectronics, Edinburgh, Scotland, UK). Automatic gain control was disabled to preserve relative signal levels. The codec's highpass filter setting was modified to a cutoff frequency of 10 Hz.

Alignment of the MIC and ACC signal was achieved using a custom algorithm in MATLAB (The MathWorks, Natick, MA) that shifted the ACC signal (upsampled to the acoustic sampling rate of 20 kHz) such that the absolute value of the cross-correlation between the two signals was maximized. This alignment inherently compensated for time delays associated with the acoustic propagation time between ACC and MIC sensors. The middle 0.5 s was extracted from each vowel waveform for vocal function analysis to capture a quasi-steady state segment unaffected by transient onset and offset behaviors.

### C. Vocal function measures

Vocal function measures consisted of three types: (1) time-domain perturbation (jitter, shimmer, harmonics-to-noise ratio), (2) spectral harmonicity (harmonics-to-noise ratio, spectral tilt), and (3) cepstral periodicity (cepstral peak prominence).

**1) Time-domain perturbation**—Time-domain perturbation measures of jitter, shimmer, and harmonics-to-noise ratio were computed to capture time-varying features of period duration and amplitude that are hypothesized to translate from the ACC to MIC waveforms.

Figure 2A illustrates the identification of glottal pulse timings  $t_i$  and amplitudes  $a_i$  for cycle index  $i$  by Praat on example waveforms [34]. Praat estimates the  $f_0$  contour using a time-domain autocorrelation approach, followed by the assignment of a time instant at each pulse at the maximum absolute amplitude within each period. In certain cases, the start-up processes in the algorithm yielded glottal pulse timings that were offset in the MIC signal relative to the ACC signal (38.2% of recordings). In those cases, the pairing of glottal cycles in the acoustic and acceleration waveforms was compensated for by a custom MATLAB algorithm that shifted the glottal pulse instants such that the cycle-to-cycle  $f_0$  (reciprocal of each period) maximally correlated between acoustic and acceleration domains for each recording. This post-alignment timing compensation is an important step for any algorithm analyzing pairs of independent signals, even when applied to quasi-stationary signals.

Figure 3 shows the effect of the glottal pulse timing compensation by displaying scatter plots of the  $f_0$  of periods within one recording of a female subject. The correlation increases from 0.83 to 0.98, demonstrating that pulse-to-pulse alignment is as accurate as possible after glottal pulse timing compensation. Glottal cycle period durations were defined as  $p_i = t_{i+1} -$

$t_i$  with average period duration  $\bar{p} = \frac{1}{N-1} \sum_{i=1}^{N-1} p_i$ , where  $N$  is the number of glottal pulse timings found.

In the voice literature, many variants of jitter and shimmer have been proposed and studied in the area of acoustic voice analysis [35]. These perturbation algorithms typically employ some type of stability analysis to determine the degree of perturbation of a time series (cycle-to-cycle period duration and amplitude for jitter and shimmer, respectively). Parameters of these perturbation analyses include definitions of short-term variability, the use of the derivative operator, and smoothing factors.

For the current study, two variants of jitter and shimmer were computed using  $p_i$  and corresponding amplitudes  $a_p$ , respectively. The first variant, the coefficient of variation (standard deviation divided by the mean) of the period durations was defined as

$$J_{CV} \triangleq \frac{1}{\bar{p}} \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (p_i - \bar{p})^2}. \quad (1)$$

$J_{CV}$  was selected as a measure of overall average variability of a time series with a stable mean.

To quantify cycle-to-cycle changes in period duration that take into account time ordering within the time series, local jitter [36], [37] was defined as the average absolute difference between consecutive period durations divided by the mean period duration:

$$J_{local} \triangleq \frac{1}{\bar{p}(N-2)} \sum_{i=1}^{N-2} |p_{i+1} - p_i|. \quad (2)$$

Although many other definitions of jitter exist [e.g., smoothing parameters or temporal units used [35]], the two variants implemented provided initial results regarding the ability of global variability (no time ordering) and cycle-to-cycle variability (time ordering) to be mapped from the MIC signal to the ACC signal.

The two variants of shimmer— $S_{CV}$  and  $S_{local}$ —were computed analogously:

$$S_{CV} \triangleq \frac{1}{\bar{a}} \sqrt{\frac{1}{N} \sum_{i=1}^N (a_i - \bar{a})^2} \quad (3)$$

and

$$S_{local} \triangleq \frac{1}{\bar{a}(N-1)} \sum_{i=1}^{N-1} |a_{i+1} - a_i|, \quad (4)$$

where the average glottal cycle instantaneous amplitude  $\bar{a} = \frac{1}{N} \sum_{i=1}^N a_i$ .



A time-domain estimate of harmonics-to-noise ratio ( $\text{HNR}_{\text{time}}$ ) was computed using Praat through the *Harmonicity (cross-correlation method)* function [37]. The method employs template matching over sliding windows in the waveform, with variations in waveshape represented as the noise component.

**2) Spectral domain**—Fig. 2B illustrates a representative average magnitude spectrum of the MIC and ACC waveforms. A spectral measure of harmonics-to-noise ratio ( $\text{HNR}_{\text{spec}}$ ) was performed using a periodic/noise decomposition method that employs a comb filter to extract the harmonic component of a signal [38], [39]. This ‘pitch-scaled harmonic filter’ approach uses an analysis window duration equal to an integer number (four in the current work) of local periods and relies on the property that harmonics of the  $f_0$  exist at specific frequency bins (every fourth bin) of the discrete-time Fourier transform. The harmonic component of the signal was thus estimated from the comb-filtered spectrum. Subtraction of the harmonic spectrum from the original waveform’s spectrum yielded the noise component, where spectral interpolation filled in gaps in the residual noise spectrum formerly containing harmonic power. The inverse discrete-time Fourier transform of the harmonic and noise spectral estimates yielded time-domain waveforms that are joined via overlap-add synthesis of successive analysis windows.  $\text{HNR}_{\text{spec}}$  was the ratio, in dB, of the power of the decomposed harmonic and noise signals.

A measure of spectral tilt ( $\text{TL}_g$ ) was implemented to provide an estimate of the spectral slope over the first 8 harmonics associated with the voice source [40]. Each waveform was divided into eight analysis frames (Hamming windowed) with 50 % overlap, and an average spectrum obtained over the analysis frames (modified periodogram via Welch’s method). Harmonic amplitudes were estimated from the average spectrum as the first 8 peaks in the vicinity ( $\pm 50$  Hz) of integer multiples of the mean fundamental frequency of the waveform. For the MIC signals, harmonic amplitudes were compensated for the amplifying effects of the first three formants [41] to yield a voice source–related decay of harmonic amplitudes. Mean formant frequencies and bandwidths were estimated by Praat using the *Burg* method across 50 ms analysis frames with 50 % overlap [37]. Finally,  $\text{TL}_g$  was computed as the linear regression slope, in dB per octave, over the first 8 compensated harmonic magnitudes. Analysis of the ACC waveforms did not include formant compensation because minimal residual formant information was present in the ACC signal; subglottal resonances information in the ACC signal was unfiltered.

**3) Cepstral domain**—Fig. 2C displays an overlay of example cepstra of the MIC and ACC waveforms. Recently, measures derived from the acoustic cepstrum have been adapted for clinical voice assessment using the commercially available Analysis of Dysphonia in Speech and Voice (ADSV) program (PENTAX, Lincoln Park, NJ). Since the cepstrum can be calculated in many different ways, a brief description of the method used by the ADSV program is given here [42]. After resampling the waveform to 25 kHz, the signal was Hamming-windowed into 40.96 ms (1024 samples) frames with 75% overlap. Two 1024-point discrete Fourier transforms were computed in succession with a logarithmic transformation between them. A 7-frame cepstral smoothing was performed by averaging the power cepstrum with those of the three frames prior to and following a given frame. Due

to a bias in the power cepstrum, a regression line was computed over frequencies greater than 2 ms (corresponding to a frequency range minimally affected by vocal tract-related information). Finally, the cepstral peak prominence (CPP) for each analysis frame was defined as the difference, in dB, between the magnitude of the highest peak and the baseline regression level in the averaged power cepstrum. The peak search was limited to frequencies between 3.3 ms and 16.7 ms, corresponding to fundamental frequencies of 300 Hz and 60 Hz, respectively. The final CPP measure was an average over all analysis frames for each waveform.

#### D. Statistical analysis

Pearson's correlation coefficient  $r$  was used to evaluate the relationships between vocal function measures estimated from the MIC and ACC waveforms during sustained vowel production. Instantaneous glottal pulse  $f_0$  and amplitude were correlated within each recording and across all subject recordings. The set of vocal function measures was computed for each recording, and pairwise correlations between analogous algorithms on MIC and ACC signals were performed across subject recordings. Due to the computation of multiple statistical comparisons for each vowel, the baseline alpha level (0.01) of the correlation coefficients were Bonferroni corrected to mitigate the possibility of false positive results; i.e., correlation coefficients were considered statistically significant when  $p < 0.001$ .

The study also investigated any dependence of the correlations on vowel type (/a/, /i/, /u/) and any systematic bias exhibited by the ACC-based vocal function measures that would indicate under- or over-estimation of a particular measure with reference to the analogous measure derived from the MIC signal. Biases were computed as the average difference between the ACC-based measures and the analogous MIC-based measures within each subject. A nonzero bias was considered statistically significant when the associated paired t-test achieved statistical significance ( $p < 0.001$ ).

Outliers were removed prior to computing correlation coefficients if jitter/shimmer values (on either MIC or ACC signals) were greater than 5% [43] or values were three standard deviations away from the mean (8.9% outliers for shimmer metrics, 5.4% outliers for the rest of the measures). Data from the same subject on multiple visits were included in the analysis because of the general independence of the voice samples at points during the course of treatment.

All statistical analyses were performed across all subject recordings, as well as separately on the normative and patient groups.

### III. Results

Figure 4 displays scatter plots showing the co-relation of the vocal function measures for each vowel type and any systematic biases exhibited by the ACC-based measures with reference to MIC-based measures. No systematic differences were found in the ability to derive the reported vocal function measures from the ACC signal for speakers grouped by the presence or absence of a voice disorder.



## A. Correlations of vocal function measures

Table I reports the correlation coefficients between ACC-based and MIC-based measures of jitter ( $J_{CV}$ ,  $J_{local}$ ), shimmer ( $S_{CV}$ ,  $S_{local}$ ), harmonics-to-noise ratio ( $HNR_{time}$ ,  $HNR_{spec}$ ), spectral tilt ( $TL_8$ ), and cepstral peak prominence (CPP). All correlation coefficients achieved statistical significance for each subject group. The strength of the correlations varied depending on the measure computed, with correlations generally in a similar relative range within each vowel type; although the highest correlations tended to be exhibited when speakers produced the vowel /a/. The strongest correlations were obtained for the two time-domain jitter measures ( $J_{CV}$  and  $J_{local}$ ) and CPP, with  $J_{CV}$  and CPP also exhibiting the highest and most consistent correlations across different vowels.  $J_{CV}$  correlations reached as high as 0.99 for the vowel /a/, and CPP correlations peaked at 0.90 for the vowel /a/.

The ACC signal appears to capture the overall variance of period durations sensed by the MIC using the  $J_{CV}$  measure. Jitter measures depending on time order ( $J_{local}$ ) resulted in slightly lower correlations.

Overall, amplitude-based perturbation metrics of shimmer compared less well than the time-based measures of jitter between MIC and ACC signals. In the pooled group, the degree of correlation ranged from 0.44 to 0.74 for  $S_{CV}$  and from 0.24 to 0.35 for  $S_{local}$ . Spectral measures of  $HNR_{spec}$  and  $TL_8$  exhibited the lowest correlations (as high as 0.57 for  $TL_8$  on /a/) when comparing analogous measures from the MIC and ACC waveforms.

## B. Bias in accelerometer-based measures

Table II lists statistical biases observed when computing the vocal function measures from the neck-surface ACC signal in each subject group. Bias values were obtained by calculating the mean of the differences between measures derived from the ACC and MIC waveforms for each recording. Positive values indicated that the ACC-based value was higher on average with reference to the associated MIC-based value for a particular measure and vowel type.

Jitter measures demonstrated little to zero bias, whereas shimmer measures from the ACC signal tended to underestimate shimmer from the MIC signal from 0.48 to 2.00 percentage points in the pooled group. In contrast, the harmonics-to-noise ratio measures were significantly higher (up to 10.92 dB) when computed from the ACC signal, and the spectral magnitude of the first 8 ACC harmonic components decayed at a faster rate (approximately 7 dB/oct faster) than the decay rate of the first 8 MIC harmonic components. In contrast, the average CPP measure was more stable, within 1.37 dB when measured in either ACC or MIC domains.

## C. Instantaneous f0 and amplitude

Recall Fig. 3B that illustrated the ability of the instantaneous f0—reciprocal of the duration of an individual glottal cycle—to correlate highly ( $r=0.98$  in that example) when computed from both MIC and ACC domains. Pearson's correlation coefficient between MIC-based and ACC-based values of instantaneous f0 across all subject recordings was 0.99, 0.98, 0.98 for /a/, /i/, and /u/, respectively. Comparing the average f0 between ACC and MIC domains

resulted in near-perfect correlations for all vowel types. Pitch halving errors were experienced in a few cases when analyzing the MIC waveform, whereas corresponding processing of the ACC waveform was unaffected by these errors and yielded correct  $f_0$  detection.

Analogous to the instantaneous  $f_0$  plot in Fig. 3B, comparisons were also made between the instantaneous glottal pulse amplitude (magnitude) of the same cycle in the MIC and ACC waveforms. Since signal amplitudes were in normalized units, Pearson's correlation coefficients were obtained within each subject recording only. To obtain an estimate of the variation of the amplitude correlation across subject recordings, the mean (standard deviation) of the correlation coefficients were computed for the vowels /a/, /i/, and /u/ as 0.67 (0.24), 0.58 (0.25), 0.61 (0.27), respectively. These results indicate the limited ability of the ACC signal to track cycle-to-cycle changes in amplitude that occurred in the MIC signal.

#### IV. Discussion

Given the potential advantages in clinical and ambulatory voice assessment that could be obtained by increased utilization of the subglottal neck-surface acceleration signal, this study sought to determine the extent to which this signal could be used to estimate vocal function-related parameters that are analogous to those currently obtained from the MIC signal. Models linking neck surface acceleration to radiated acoustic pressure during vowel production have employed lumped-element representations of the subglottal tract, supraglottal tract, and neck skin in an effort to derive voice source-related properties such as maximum flow declination rate and spectral tilt [44], [45], [14], [28]. The transformation from acoustics to acceleration can be represented by a linear filter such that an inverse filter may recover the glottal source waveform [14]. These results imply the opportunity for capturing information related to acoustic perturbation and spectral measures because the acoustic-to-acceleration conversion is modeled as a time-invariant transformation. These models assume that the primary contribution to the ACC waveform arises from acoustic and aerodynamic energy at the glottis. Thus the model assumptions are potentially correct in some respects (i.e., that glottal noise information is not contributing much to the acceleration waveform) and incorrect in others (i.e., vocal fold collision forces may be contributing significantly to the differences seen between the measures derived from acoustic and acceleration signals).

The moderate-to-strong correlations between ACC- and MIC-based measures of perturbation and cepstral peak prominence provide evidence that 1) clinical voice assessment could be potentially enhanced/improved by the acquisition of the subglottal neck-surface acceleration signal, particularly in noisy environments; and 2) ambulatory voice monitoring systems that use neck-placed ACCs to sense phonation could be updated to include estimates of vocal function, which could be particularly useful in increasing the versatility of those systems that also provide ambulatory biofeedback for carryover of voice therapy goals. For example, due to the high correlation between ACC- and MIC-based estimates of CPP, future work could track changes in CPP from a speaker's ambulatory accelerometer signal to reveal deterioration of vocal function over the course of a day due to vocal fatigue, etc.

ACC-based measures of harmonics-to-noise ratio and spectral tilt displayed weaker correlations with acoustic versions of these measures, indicating that these acceleration-derived measures are potentially providing different information than the acoustic measures. For example, measuring spectral tilt from the ACC signal showed a bias of approximately  $-7$  dB/oct with reference to spectral tilt computed from the acoustic waveform. This empirical value agrees well with the lowpass-filter quality of the neck frequency response that was observed to be  $-8.4$  dB/oct for individuals with normal neck tissue [28] and  $-8.8$  dB/oct across both laryngectomy patients and normal subjects [29]. Indeed, Fig. 2B illustrates that the bandwidth of the radiated acoustic signal typically surpassed 8 kHz, whereas little energy was present above 4 kHz in the ACC signal. The agreement of the ACC spectral tilt bias observed in the current study to the neck frequency response reported in the literature provide support that the spectral slope of the pneumotachograph mask was largely taken into account by the spectral compensation algorithm performed on the MIC signal prior to the computation of spectral tilt (Sec. II-C2).

The potential of subglottal neck-surface acceleration to contribute additional information regarding glottal source aerodynamics and vocal fold collision forces merits further study. It has been speculated that the neck-skin surface wave may contain contributions due to the propagation of mechanical energy of vocal fold collision forces radiating through laryngeal and neck tissue as well as to the propagation of acoustic pressure [7]. The estimation of these collision forces from the subglottal neck-surface acceleration signal may help in better understanding the pathophysiology of phonotraumatic vocal fold lesions (e.g., nodules and polyps) and hyperfunctional behaviors associated with their development [46]. In fact, when measuring vocal fold collision forces using a contact transducer, the potential confound of capturing information related to glottal air pressures on the sensor must be taken into account. Disambiguation of the various energy sources may be performed through the precise delineation of the timing of peaks in the acoustic, accelerometric, force, and electroglottographic signals within a glottal cycle [8], [47].

ACC sensors are designed to be immune to acoustic noise sources in the environment [6]. This study obtained further evidence showing that the ACC sensor also exhibited a reduced sensitivity to acoustic turbulence generated at the glottis due to the large positive bias of measuring  $HNR_{spec}$  from neck-surface acceleration relative to the acoustic pressure measure. In addition,  $HNR_{spec}$  was not highly correlated between the two domains. Although these findings might point to potential challenges in characterizing speakers with highly breathy and/or dysphonic voice qualities, acoustic CPP measures correlated highly with CPP measures computed from subglottal neck-surface acceleration. The amplitude of the first harmonic in the cepstrum is thought to act as a geometric mean HNR [31]; however, a reduced sensitivity to the turbulent noise component of the voice source forces the accelerometric CPP to reflect primarily aperiodicity of the waveform. Thus, orthogonal sets of measures may be established by accelerometric CPP for noise-free periodicity and acoustic HNR for perturbation-free turbulent noise estimation [48].

The results of this study have implications for both understanding physiological mechanisms of speech production and clinical voice assessment. In this study, no differences were found in the ability to derive the reported vocal function measures from the ACC signal for

speakers grouped by the presence or absence of a voice disorder. This result may be in part due to the fact that most patients in this study exhibited mild-to-moderate dysphonia, and therefore practically all the perturbation measures were valid. The high correlations exhibited by timing-based measures of instantaneous  $f_0$  and jitter, as well as CPP, bode well for the derivation of recently popular measures of vocal hyperfunction and dysphonia—relative fundamental frequency [11] and the cepstral/spectral index of dysphonia [42]—solely from subglottal neck-surface acceleration. Based on the encouraging result with CPP, future work calls for further investigation into vocal function assessment during continuous speech segments and on a larger sample of subjects with severely dysphonic characteristics. Furthermore, results of this study may act as a guide for applying vocal function features as inputs to machine learning algorithms to aid in the classification and/or online monitoring of voice disorders [21].

## V. Conclusion

This study provides evidence supporting the derivation of measures related to vocal function using an ACC sensor placed on the subglottal surface of the neck above the sternal notch. Strong relationships were observed between timing-based jitter of neck-surface vibration and radiated acoustic sound pressure. Lower correlations exhibited by amplitude-based measures of shimmer may point to combined effects of subglottal filtering and contributions of vocal fold impact forces. Spectral measures of harmonics-to-noise ratio (reflecting aspiration noise energy) appeared to be significantly reduced with a steeper spectral tilt (harmonic amplitude decay) when obtained from the subglottal neck-surface acceleration signal. Finally, the cepstral peak prominence correlated well between the MIC and ACC domains to reflect similar levels of periodicity in the voice source. The ability to derive vocal function measures using low-profile and confidential ACC signals has implications for noninvasive and noise-robust voice assessment, ambulatory monitoring of vocal health, and real-time biofeedback.

## Acknowledgments

This project was supported by a grant from the NIH National Institute on Deafness and Other Communication Disorders (R33 DC011588) and the Voice Health Institute. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

## References

1. Sundberg J. Chest wall vibrations in singers. *J. Speech Hear. Res.* 1983; 26(3):329–340. [PubMed: 6645456]
2. Stevens KN, Kalikow DN, Willemain TR. A miniature accelerometer for detecting glottal waveforms and nasalization. *J. Speech Hear. Res.* 1975; 18(3):594–599. [PubMed: 1186168]
3. Horii Y. An accelerometric measure as a physical correlate of perceived hypernasality in speech. *J. Speech Hear. Res.* 1983; 26(3):476–480. [PubMed: 6645475]
4. Lulich SM, Morton JR, Arsikere H, Sommers MS, Leung GKF, Alwan A. Subglottal resonances of adult male and female native speakers of american english. *J. Acoust. Soc. Am.* 2012; 132(4):2592–2602. [PubMed: 23039452]
5. Quatieri TF, Brady K, Messing D, Campbell JP, Campbell WM, Brandstein MS, Weinstein CJ, Tardelli JD, Gatewood PD. Exploiting nonacoustic sensors for speech encoding. *IEEE Trans. Audio Speech Lang. Processing.* 2006; 14(2):533–544.

6. Zañartu M, HO JC, Kraman SS, Pasterkamp H, Huber JE, Wodicka GR. Air-borne and tissue-borne sensitivities of bioacoustic sensors used on the skin surface. *IEEE Trans. Biomed. Eng.* 2009; 56(2): 443–451. [PubMed: 19272887]
7. Coleman RF. Comparison of microphone and neck-mounted accelerometer monitoring of the performing voice. *J. Voice.* 1988; 2(3):200–205.
8. Gunter HE, Howe RD, Zeitels SM, Kobler JB, Hillman RE. Measurement of vocal fold collision forces during phonation: Methods and preliminary data. *J. Speech. Lang. Hear. Res.* 48(3):567–576. 2005. [PubMed: 16197273]
9. Sugimoto T, Hiki S. Extraction of the pitch of a voice from the vibration of the outer skin of the trachea. *Journal of the Acoustical Society of Japan.* 1960; 1(4):291–293.
10. Porter HC. Extraction of pitch from the trachea. *Air Force Cambridge Research Laboratories Research Note.* 1963; 63–24
11. Lien Y-AS, Stepp CE. Comparison of voice relative fundamental frequency estimates derived from an accelerometer signal and low-pass filtered and unprocessed microphone signals. *J. Acoust. Soc. Am.* 2014; 135(5):2977–2985. [PMID: 24815277]. [PubMed: 24815277]
12. Švec JG, Titze IR, Popolo PS. Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *J. Acoust. Soc. Am.* 2005; 117(3):1386–1394. [PubMed: 15807026]
13. Hillman RE, Heaton JT, Masaki A, Zeitels SM, Cheyne HA. Ambulatory monitoring of disordered voices. *Ann. Otol. Rhinol. Laryngol.* 2006; 115(11):795–801. [PubMed: 17165660]
14. Zañartu M, Ho JC, Mehta DD, Hillman RE, Wodicka GR. Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Trans. Audio Speech Lang. Processing.* 2013; 21(9):1929–1939.
15. Askenfelt A, Gauffin J, Sundberg J. A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. *J. Speech Hear. Res.* 1980; 23(2):258–273. [PubMed: 7003261]
16. Szabo A, Hammarberg B, Hkansson A, Sdersten M. A voice accumulator device: Evaluation based on studio and field recordings. *Logopedics, Phoniatrics, Vocology.* 2001; 26(3):102–117.
17. Horii Y. Jitter and shimmer differences among sustained vowel phonations. *J. Speech Hear. Res.* 1982; 25(1):12–14. [PubMed: 7087413]
18. Buekers R, Bierens E, Kingma H, Marres E. Vocal load as measured by the voice accumulator. *Folia Phoniatr. Logop.* 1995; 47(5):252–261. [PubMed: 8563777]
19. Titze IR, Švec JG, Popolo PS. Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *J. Speech. Lang. Hear. Res.* 2003; 46(4):919–932. [PubMed: 12959470]
20. Mehta DD, Zañartu M, Feng SW, Cheyne II HA, Hillman RE. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Trans. Biomed. Eng.* 2012; 59(11):3090–3096. [PubMed: 22875236]
21. Ghassemi M, Van Stan JH, Mehta DD, Zañartu M, Cheyne HA II, Hillman RE, Guttag JV. Learning to detect vocal hyperfunction from ambulatory neck-surface acceleration features: Initial results for vocal fold nodules. *IEEE Trans. Biomed. Eng.* 2014; 61(6):1668–1675. [PubMed: 24845276]
22. Mehta DD, Van Stan JH, Zañartu M, Ghassemi M, Guttag JV, Espinoza VM, Cortés JP, Cheyne HA, Hillman RE. Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Front. Bioeng. Biotechnol.* 2015; 3(155):1–14. in press [Online]. Available: [http://www.frontiersin.org/Journal/Abstract.aspx?s=1267&name=bioinformatics\\_and\\_computational\\_biology&ART\\_DOI=10.3389/fbioe.2015.00155](http://www.frontiersin.org/Journal/Abstract.aspx?s=1267&name=bioinformatics_and_computational_biology&ART_DOI=10.3389/fbioe.2015.00155). [PubMed: 25654078]
23. Lieberman P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.* 1963; 35(3):344–353.
24. Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.* 1990; 87(2):820–857. [PubMed: 2137837]
25. Mehta DD, Zeitels SM, Burns JA, Friedman AD, Deliyski DD, Hillman RE. High-speed videoendoscopic analysis of relationships between cepstral-based acoustic measures and voice production mechanisms in patients undergoing phonemicsurgery. *Ann. Otol. Rhinol. Laryngol.* 2012; 121(5):341–347. [PubMed: 22724281]

26. Roy N, Barkmeier-Kraemer J, Eadie T, Sivasankar MP, Mehta D, Paul D, Hillman R. Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*. 2013; 22(2):212–226.
27. Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice*. 2010; 24(5):540–555. [PubMed: 19883993]
28. Wu L, Xiao K, Dong J, Wang S, Wan M. Measurement of the sound transmission characteristics of normal neck tissue using a reflectionless uniform tube. *J. Acoust. Soc. Am*. 2014; 136(1):350–356. [PubMed: 24993219]
29. Meltzner GS, Kobler JB, Hillman RE. Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *J. Acoust. Soc. Am*. 2003; 114(2):1035–1047. [PubMed: 12942982]
30. Fraile R, Godino-Llorente JI. Cepstral peak prominence: A comprehensive analysis. *Biomed. Signal Process. Control*. 2014; 14:42–54.
31. Murphy PJ. On first harmonic amplitude in the analysis of synthesized aperiodic voice signals. *J. Acoust. Soc. Am*. 2006; 120(5):2896–2907. [PubMed: 17139747]
32. Murphy PJ, Akande OO. Noise estimation in voice signals using short-term cepstral analysis. *J. Acoust. Soc. Am*. 2007; 121(3):1679–1690. [PubMed: 17407904]
33. Mehta, DD.; Zaartu, M.; Van Stan, JH.; Feng, SW.; Cheyne II, HA.; Hillman, RE. Smartphone-based detection of voice disorders by long-term monitoring of neck acceleration features; Proceedings of the 10th Annual Body Sensor Networks Conference; 2013.
34. Boersma, P. Institute of Phonetic Sciences. University of Amsterdam; 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,” in; p. 97-110.
35. Buder, EH. Acoustic analysis of voice quality: A tabulation of algorithms 1902–1990. In: Kent, RD.; Ball, MJ., editors. *Voice Quality Measurement*. San Diego, CA: Singular Thomson Learning; 2000. p. 119-244.
36. KayPENTAX. Software instruction manual: Multi-dimensional voice program (mdvp) model 5105. 2007
37. Boersma, P.; Weenink, D. [last viewed 13 July 2009] Praat: Doing phonetics by computer. version 5.1.40. University of Amsterdam, The Netherlands. 2009. <http://www.fon.hum.uva.nl/praat/>
38. Jackson PJB, Shadle CH. Frication noise modulated by voicing, as revealed by pitch-scaled decomposition. *J. Acoust. Soc. Am*. 2000; 108(4):1421–1434. [PubMed: 11051468]
39. Jackson PJB, Shadle CH. Pitch-scaled estimation of simultaneous voiced and turbulencenoise components in speech. *IEEE Trans. Speech Audio Process*. 2001; 9(7):713–726.
40. Mehta DD, Zaňartu M, Quatieri TF, Deliyski DD, Hillman RE. Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videendoscopy. *J. Acoust. Soc. Am*. 2011; 130(6):3999–4009. [PubMed: 22225054]
41. Iseli M, Shue Y-L, Alwan A. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *J. Acoust. Soc. Am*. 2007; 121(4):2283–2295. [PubMed: 17471742]
42. Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clin. Linguist. Phon*. 2010; 24(9):742–758. [PubMed: 20687828]
43. Titze IR. Workshop on acoustic voice analysis: Summary statement,”. National Center for Voice and Speech, Tech. Rep. 1995 Feb 17–18.:1994.
44. Cheyne, HA. Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck; Proceedings of the 3rd IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors; 2006. p. 118-121.
45. Lulich SM, Alwan A, Arsikere H, Morton JR, Sommers MS. Resonances and wave propagation velocity in the subglottal airways. *J. Acoust. Soc. Am*. 2011; 130(4):2108–2115. [PubMed: 21973365]
46. Hillman RE, Holmberg EB, Perkell JS, Walsh M, Vaughan C. Objective assessment of vocal hyperfunction: An experimental framework and initial results. *J. Speech Hear. Res*. 1989; 32(2): 373–392. [PubMed: 2739390]



47. Li Z, Bakhshae H, Helou L, Mongeau L, Kost K, Rosen C, Verdolini K. Evaluation of contact pressure in human vocal folds during phonation using high-speed videoendoscopy, electroglottography, and magnetic resonance imaging. *Proceedings of Meetings on Acoustics*. 2013; 19(1):06036. [Online]. Available: <http://scitation.aip.org/content/asa/journal/poma/19/1/10.1121/1.4800732>.
48. Murphy PJ. Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis. *J. Acoust. Soc. Am.* 1999; 105(5):2866–2881. [PubMed: 10335636]

## Biographies



**Daryush D. Mehta** (S'01–M'11) received the B.S. degree in electrical engineering (summa cum laude) from University of Florida, Gainesville, in 2003, the S.M. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 2006, and the Ph.D. degree from MIT in speech and hearing bioscience and technology in the Harvard–MIT Division of Health Sciences and Technology, Cambridge, in 2010.

He currently holds appointments at Massachusetts General Hospital (Assistant Biomedical Engineer in the Department of Surgery), Harvard Medical School (Instructor in Surgery), and the MGH Institute of Health Professions (Adjunct Assistant Professor in the Department of Communication Sciences and Disorders), Boston. He is also an Honorary Senior Fellow in the Department of Otolaryngology, University of Melbourne, in Australia.



**Jarrad H. Van Stan** Jarrad H. Van Stan received the B.M. degree in applied voice from the University of Delaware, Newark, USA, in 2001, and the M.A. degree in speech pathology from Temple University, Philadelphia, PA, USA, in 2005.

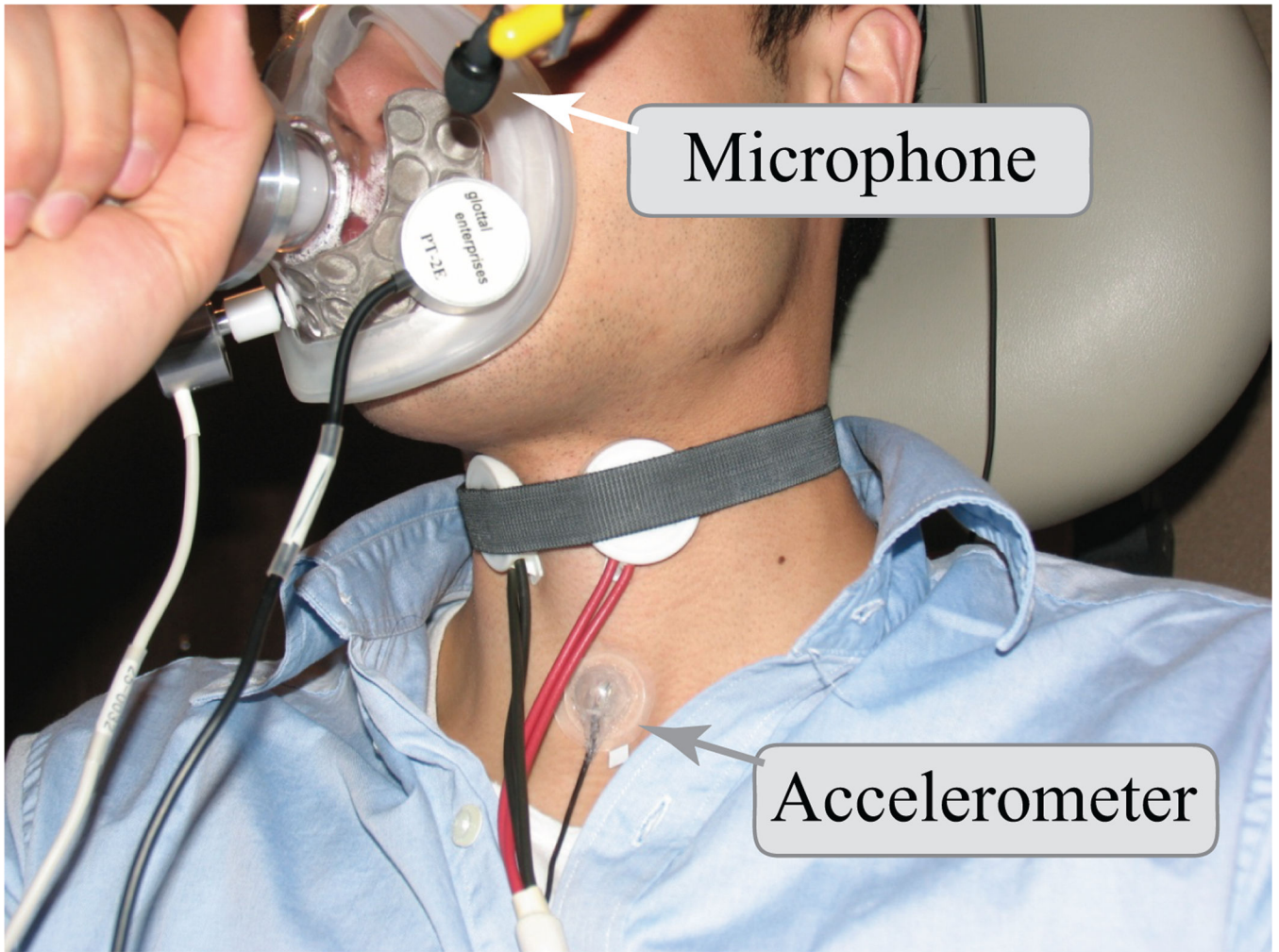
He is currently a Speech-Language Pathologist and a Senior Clinical Research Coordinator at the MGH Center for Laryngeal Surgery and Voice Rehabilitation and a Ph.D. student at the MGH Institute of Health Professions, Boston, MA, USA. He is a Board Recognized Specialist in swallowing disorders and his research interests include voice and swallowing assessment and rehabilitation.



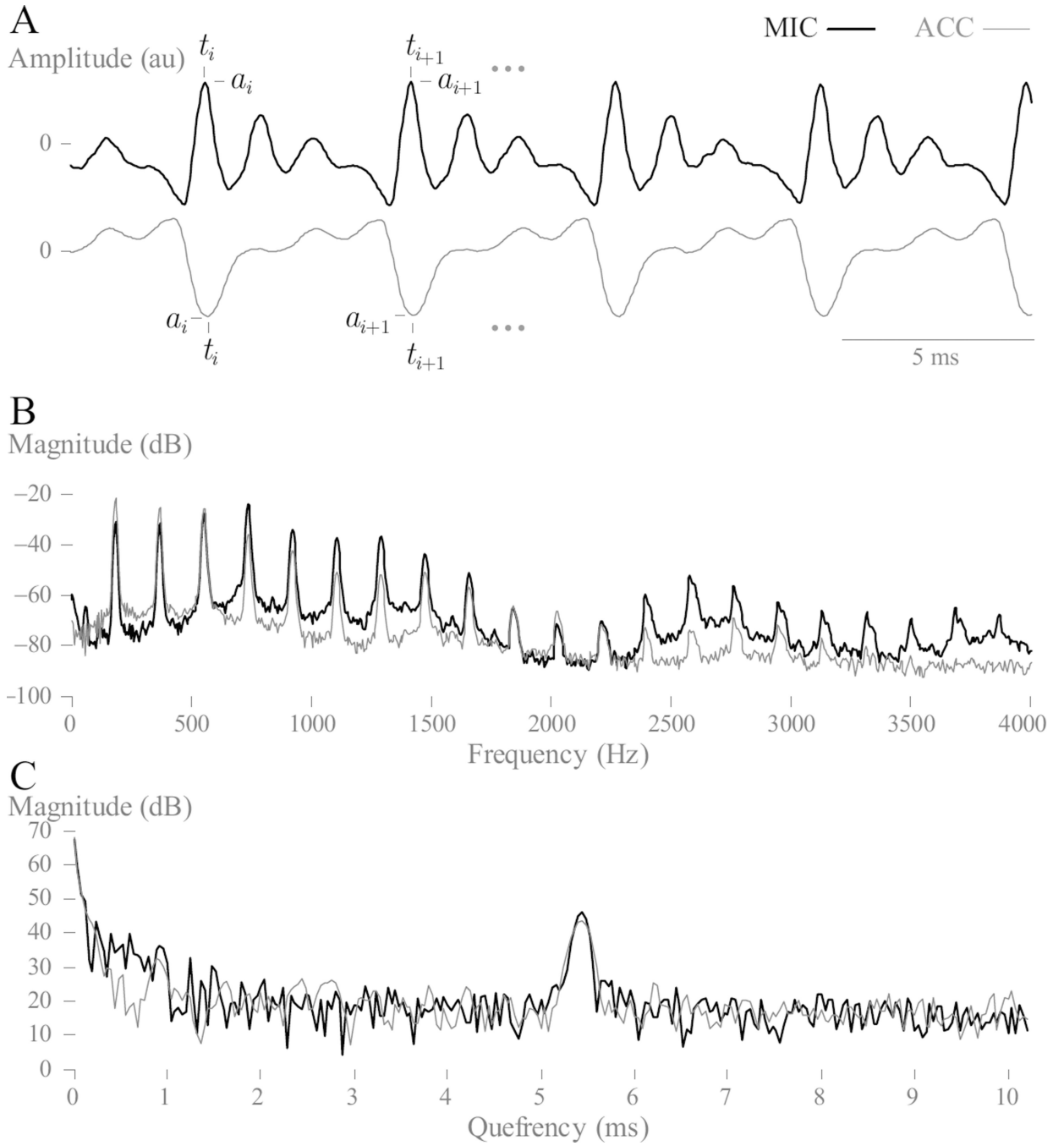
**Robert E. Hillman** received the B.S. and M.S. degrees in speech pathology from Pennsylvania State University, University Park, in 1974 and 1975, respectively, and the Ph.D. degree in speech science from Purdue University, West Lafayette, IN, in 1980.

He is currently Co-Director/Research Director of the MGH Center for Laryngeal Surgery and Voice Rehabilitation, Associate Professor of Surgery & Health Sciences and Technology at Harvard Medical School, and Professor and Associate Provost for Research at the MGH Institute of Health Professions, Boston, MA. His research has been funded by both governmental and private agencies since 1981, and he has over 100 publications on normal and disordered voice.

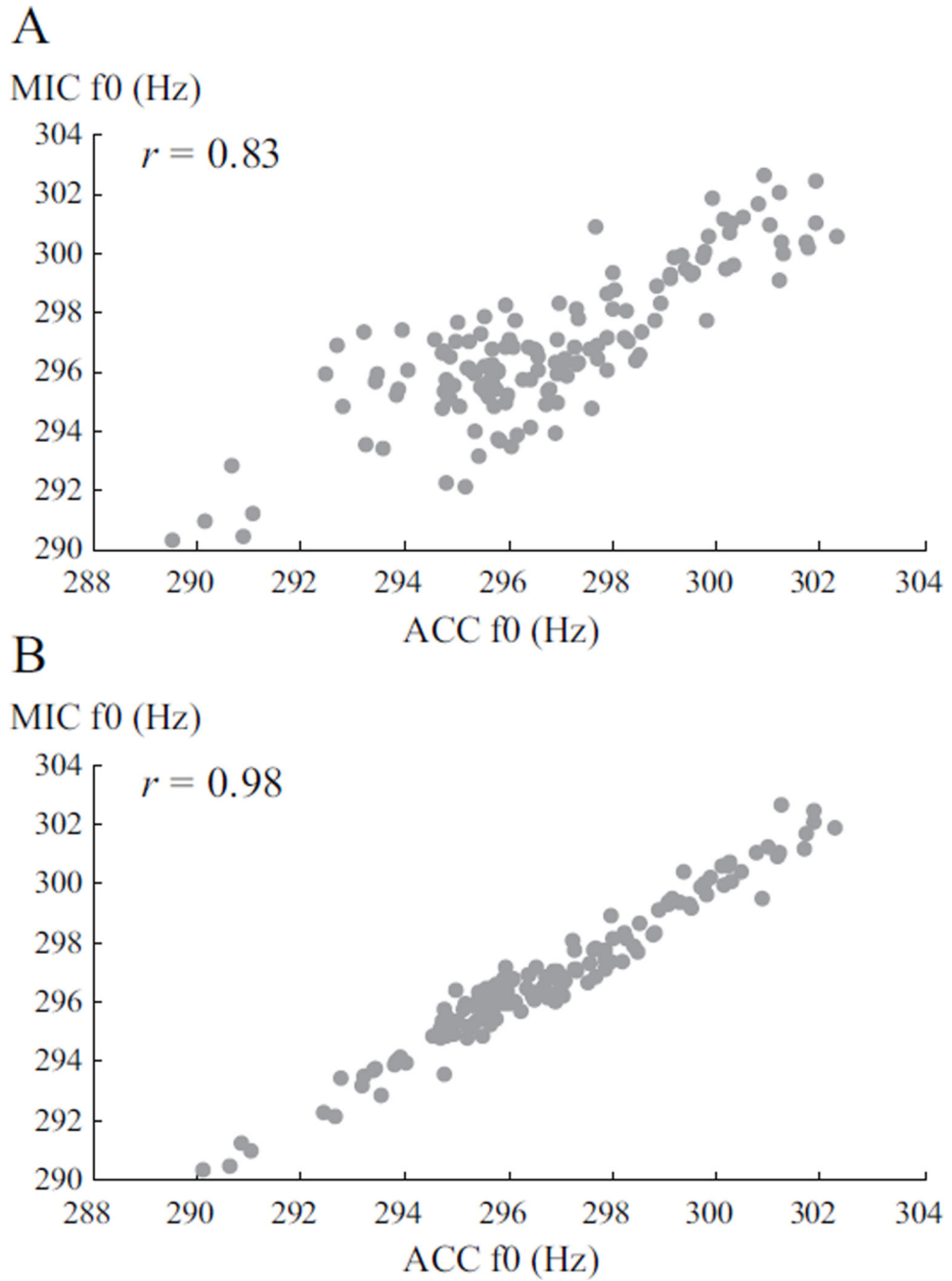
He is a Fellow of the American Speech-Language-Hearing Association (also receiving Honors of the Association, ASHA's highest honor) and the American Laryngological Association.



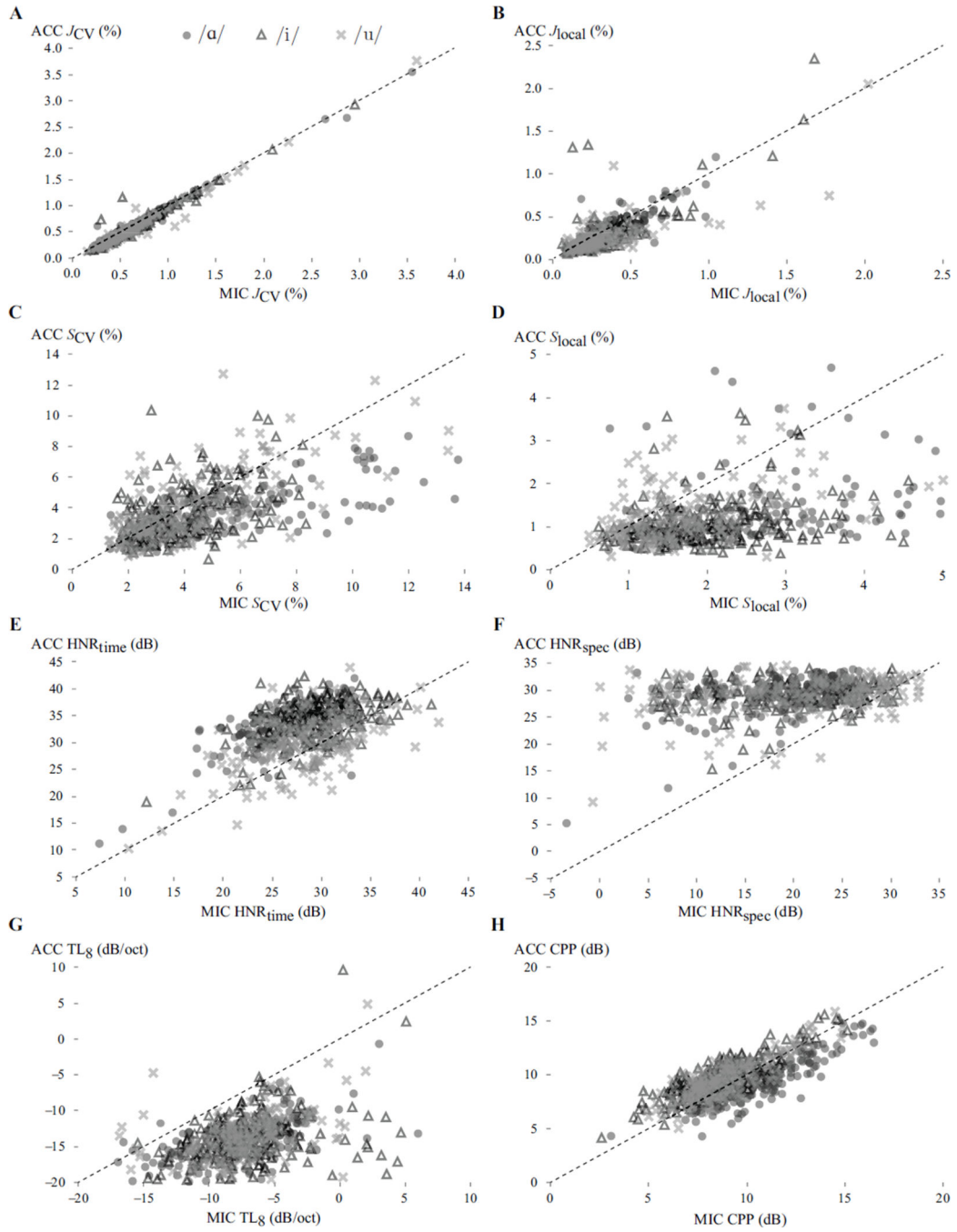
**Fig. 1.**  
(Color online) Illustration of the positioning of microphone and accelerometer sensors on a subject.



**Fig. 2.** Exemplary analysis of microphone (MIC, black) and accelerometer (ACC, gray) signals recorded during the production of the sustained vowel /a/. Snapshots are shown of the (A) time-domain waveform in linear arbitrary units (au), (B) frequency-domain average magnitude spectrum, and (C) power cepstrum. Glottal cycle indices  $i$  are identified and labeled by timing  $t_i$  and amplitude  $a_i$  parameters.



**Fig. 3.** Effect of shifting identified glottal pulse timings by one period between the microphone (MIC) and accelerometer (ACC) waveforms for the /a/ vowel produced by a vocally healthy adult female. Scatter plots and Pearson's  $r$  show the relationships between the fundamental frequency ( $f_0$ ) of each glottal cycle derived from the MIC and ACC signals (A) before and (B) after glottal pulse timing compensation.



**Fig. 4.** Comparison of vocal function measures obtained from acoustic microphone (MIC) and neck-surface accelerometer (ACC) waveforms across all 191 subject recordings for each vowel type: /a/ (filled circles), /i/ (triangles), /u/ (x's). Scatter plots are shown for the (A) coefficient of variation of period durations ( $J_{CV}$ ), (B) average first-difference of period durations ( $J_{local}$ ), (C) coefficient of variation of glottal pulse amplitudes ( $S_{CV}$ ), (D) average first-difference of glottal pulse amplitudes ( $S_{local}$ ), (E) time-domain harmonics-to-noise ratio ( $HNR_{time}$ ), (F) spectral-domain harmonics-to-noise ratio ( $HNR_{spec}$ ), (G) harmonic spectral

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



tilt ( $TL_8$ ), and (H) cepstral peak prominence (CPP). A diagonal line (dashed) through the origin with slope of 1 aids in visualizing any biases.

**TABLE I**

Pearson's correlation coefficient ( $p < 0.001$ ) for pairwise relationships between vocal function measures estimated from acoustics and neck-surface acceleration in the normal (NI), patient (Pt), and pooled (All) groups.

Measure	/a/			/i/			/u/		
	NI	Pt	All	NI	Pt	All	NI	Pt	All
$J_{CV}$	0.99	0.99	0.99	0.99	0.97	0.98	0.96	0.96	0.99
$J_{local}$	0.91	0.82	0.84	0.89	0.78	0.79	0.78	0.82	0.80
$S_{CV}$	0.65	0.79	0.74	0.40	0.47	0.44	0.55	0.72	0.69
$S_{local}$	0.25	0.35	0.33	0.15	0.27	0.24	0.22	0.39	0.35
$HNR_{time}$	0.65	0.70	0.69	0.27	0.61	0.55	0.34	0.56	0.52
$HNR_{spec}$	0.48	0.40	0.40	0.28	0.19	0.21	0.14	0.29	0.26
$TL_8$	0.52	0.59	0.57	0.05	0.30	0.23	0.45	0.46	0.46
CPP	0.90	0.90	0.90	0.90	0.87	0.88	0.89	0.82	0.84

TABLE II

Bias of computing vocal function measures using neck-surface acceleration with acoustic measure as reference ( $p < 0.001$ ) in the normal (NI), patient (Pt), and pooled (All) groups.

Measure (Units)	/a/			/i/			/u/		
	NI	Pt	All	NI	Pt	All	NI	Pt	All
$J_{CV}$ (pp)	—	—	—	-0.02	—	—	—	—	—
$J_{local}$ (pp)	—	-0.03	-0.03	-0.05	—	—	—	-0.04	-0.05
$S_{CV}$ (pp)	-2.05	-1.98	-2.00	—	-1.87	-0.48	—	—	—
$S_{local}$ (pp)	-1.29	-1.04	-1.11	-1.06	-0.95	-0.98	-0.48	-0.49	-0.49
$HNR_{time}$ (dB)	7.30	7.20	7.23	4.31	—	4.53	—	1.42	1.36
$HNR_{spec}$ (dB)	11.18	10.82	10.92	9.79	9.90	9.87	9.74	9.12	9.29
$TL_8$ (dB/oct)	-6.90	-6.66	-6.72	-7.10	-6.83	-6.91	-6.69	-6.81	-6.77
CPP (dB)	-1.51	-1.32	-1.37	1.33	1.38	1.36	0.58	0.87	0.79

pp = percentage points, dash (—) = zero bias