# Smartphone-based detection of voice disorders by long-term monitoring of neck acceleration features


Daryush D. Mehta*†, Matías Zañartu‡, Jarrad H. Van Stan*§,
Shengran W. Feng*¶, Harold A. Cheyne II‖, Robert E. Hillman*†§

*Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital, Boston, Massachusetts 02114
†Department of Surgery, Harvard Medical School, Boston, Massachusetts 02115
‡Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
§Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, Massachusetts 02129
¶Harvard-MIT Division of Health Sciences & Technology, Massachusetts Institute of Technology, Cambridge, MA 02139
‖Bioacoustic Research Program, Lab of Ornithology, Cornell University, Ithaca, New York 14853

*daryush.mehta@alum.mit.edu, matias.zanartu@usm.cl, jvanstan@mghihp.edu*
*willfeng@mit.edu, hac68@cornell.edu, hillman.robert@mgh.harvard.edu*



***Abstract*—Many common voice disorders are chronic or recurring conditions that are likely to result from inefficient and/or abusive patterns of vocal behavior, termed vocal hyperfunction. Thus an ongoing goal in clinical voice assessment is the long-term monitoring of noninvasively derived measures to track hyperfunction. This paper reports on a smartphone-based voice health monitor that records the high-bandwidth accelerometer signal from the neck skin above the collarbone. Data collection is under way from patients with vocal hyperfunction and matched-control subjects to create a dataset designed to identify the best set of diagnostic measures for hyperfunctional patterns of vocal behavior. Vocal status is tracked from neck acceleration using previously-developed vocal dose measures and novel model-based features of glottal airflow estimates. Clinically, the treatment of hyperfunctional disorders would be greatly enhanced by the ability to unobtrusively monitor and quantify detrimental behaviors and, ultimately, to provide real-time biofeedback that could facilitate healthier voice use.**



***Index Terms*—voice use, vocal hyperfunction, voice production model, accelerometer sensor, wearable voice sensor**


## I. Introduction

Voice disorders affect approximately 6.6 % of the adult population in the United States at any given point in time [1]. While most normal speakers take voice production for granted, reduction or loss of the ability to produce voice can disrupt or preclude normal oral communication and thus have far-reaching social, professional, and personal consequences [2]. Many common voice disorders are chronic or recurring conditions that are likely to result from faulty and/or abusive patterns of vocal behavior referred to generically as *vocal hyperfunction* [3]. These behaviorally based voice disorders can be especially difficult to assess accurately in the clinical setting and potentially could be much better characterized by long-term ambulatory voice monitoring as individuals engage in their typical daily activities.

### A. Behaviorally based voice disorders

A major ongoing goal of our work is the development and use of noninvasive measures to quantify vocal hyperfunction and better define its role in common voice disorders so that new clinical tools can be created to improve their diagnosis and treatment. Beginning in the late 1980s, our group proposed and began to provide evidence for the concept that there are two types of vocal hyperfunction that can be quantitatively described and differentiated from each other and normal voice production using a combination of acoustic and aerodynamic measures [3], [4]. We refer to these two types as *adducted hyperfunction* and *non-adducted hyperfunction*.

Figure 1A shows an endoscopic image of a healthy larynx with smooth and straight vocal fold edges and minimal constriction above the level of the vocal folds (Figure 1D illustrates subject endoscopy). In contrast, adducted hyperfunction is associated with the formation of benign vocal fold lesions—such as nodules and polyps—and is accompanied by abnormalities in vocal function measures. Figure 1B shows a subject with a left vocal fold polyp that is caused by vocal fold trauma due to elevated vocal fold closure velocities and tissue collision forces.

Non-adducted hyperfunction is associated with vocal fatigue and dysphonia, but with an absence of vocal fold tissue trauma. The subject imaged in Figure 1C exhibits such a non-adducted hyperfunction accompanied by excessive neck muscle tension that is typically associated with airflow parameters that reflect less complete glottal closure. Other clinical terms for non-adducted hyperfunction encountered in the literature include functional dysphonia and muscle tension dysphonia.

For the disorders linked to adducted hyperfunction, clinicians spend a significant amount of time in trying to identify and modify the prevalence of what are believed to be abusive or "phono-traumatic" vocal behaviors (e.g., excessive loudness) in daily voice. The nature and severity of the vocal behaviors associated with non-adducted hyperfunction can display significant situational variation (e.g., with levels of emotional stress) throughout the course of a day. Clinicians currently rely on patients to assess the prevalence and persistence of vocal patterns behaviors by self-reporting their own


978-1-4799-0330-6/13/$31.00 ©2013 IEEE

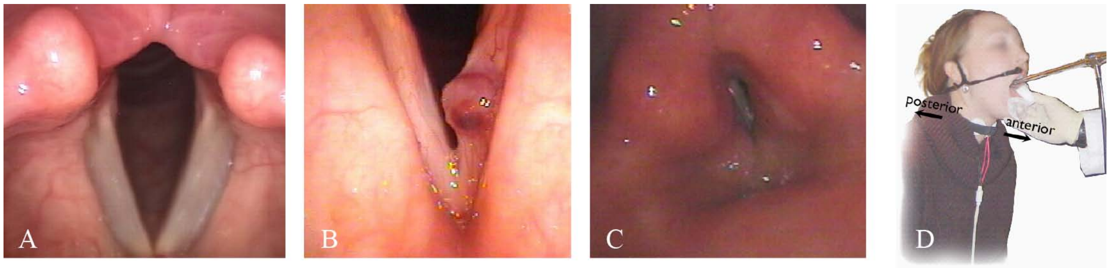


Fig. 1. Endoscopic images of the larynx of (A) a normal subject and subjects with (B) adducted hyperfunction (left vocal fold polyp) and (C) non-adducted hyperfunction (excessive constriction above the vocal folds). Left and right are reversed due to the orientation of the observer and the subject during (D) endoscopy.

voice use, which is highly subjective and prone to unreliability.

### B. *Development of a new voice health monitor*

We believe that the diagnosis and treatment of many common voice disorders would be greatly enhanced by the ability to unobtrusively monitor and quantify hyperfunctional vocal behaviors as individuals go about their normal daily activities. It is expected that this type of ambulatory monitoring would enable clinicians to better assess the role of phonatory behaviors in what are commonly viewed as voice use–related disorders, precisely pinpoint the location and duration of abusive and/or maladaptive behaviors so these can be specifically targeted for modification, and objectively assess patient compliance with the behavioral goals of voice therapy. Our initial efforts to develop an ambulatory voice monitoring system [5], [6] were largely motivated by the difficulties that are encountered in accurately assessing hyperfunctional voice disorders in the clinical setting.

Most recently we have reported on our development of a user friendly and flexible platform for ambulatory voice monitoring that uses a neck-placed miniature accelerometer as voice sensor and a smartphone as the data acquisition platform [7]. The main advantage of this device over current-generation systems is the collection of the unprocessed accelerometer signal that allows for the investigation of new voice use–related measures based on a vocal system model. In addition, the ease of use of the smartphone-based platform enables large-sample clinical studies that that can identify measures of high statistical power for differentiating between hyperfunctional and normal patterns of vocal behavior.

## II. Smartphone-based voice health monitoring

Figure 2 illustrates the coupling of the smartphone to the accelerometer sensor using an interface circuit. Minimum device recording specifications to record the raw accelerometer signal include an 11 025 Hz sampling rate, 16-bit quantization, 50-dB dynamic range, and 9.5 GB of nonvolatile memory. These specifications satisfy the requirements of obtaining voice-related neck skin vibrations from quiet-to-shouting voiced sounds with frequencies up to 4000 Hz. We selected the Nexus S smartphone manufactured by Samsung with Google's Android operating system. In addition to routines associated with controlling the data acquisition process, the programming environment enables the development of interactive applications for gathering subject self-report data, prompting daily calibration protocols, and verifying signal integrity.

The Nexus S smartphone contains a high-fidelity audio codec (WM8994; Wolfson Microelectronics, Edinburgh, Scotland, UK) that encodes the accelerometer signal using sigma-delta modulation (128x oversampling). Sampling rates available include standard frequencies from 8000 Hz to 48 kHz. The Nexus S smartphone enables one mono (microphone) input of the two stereo channels available on the codec. Of critical importance, operating system root access allows control over driver settings related to highpass filters (to remove dc offsets and low-frequency noise) and programmable gain arrays (to maximize dynamic range).

The smartphone provides power to and acquires the accelerometer signal using the microphone channel. The accelerometer requires a minimum of 1.5 V, which is satisfied by the 2 V bias signal on the microphone channel. The interface circuit provides appropriate impedance matching such that the accelerometer acts as a handsfree microphone. Although signal bandwidth and power requirements currently warrant a wired setup, future device designs could consider wireless interfaces with smaller footprints. The circuit couples the accelerometer pinout (power, ground, signal) to the smartphone microphone channel (signal+bias, ground) and is encased in an epoxy with appropriate strain relief.

### A. *In-laboratory calibration*

The vocal system model that is used to extract measures of glottal airflow from the accelerometer signal requires a subject-specific calibration that characterizes subglottal impedance and mechanical skin properties [8]. Figure 3 illustrates the setup of this one-time calibration in the laboratory that includes simultaneous recordings of neck skin acceleration, the acoustic voice signal, high-bandwidth oral airflow, oral air pressure, and electroglottography. A circumferentially-vented pneumotachograph mask system (model MA-1L; Glottal Enterprises, Syracuse, NY) yields the oral airflow and air pressure signals.

The in-laboratory calibration protocol prompts subjects with vocal hyperfunction to produce the following sounds at a comfortable pitch:

1) /a: i: u:/ at comfortable, soft, and loud levels
2) Rainbow Passage at comfortable loudness level

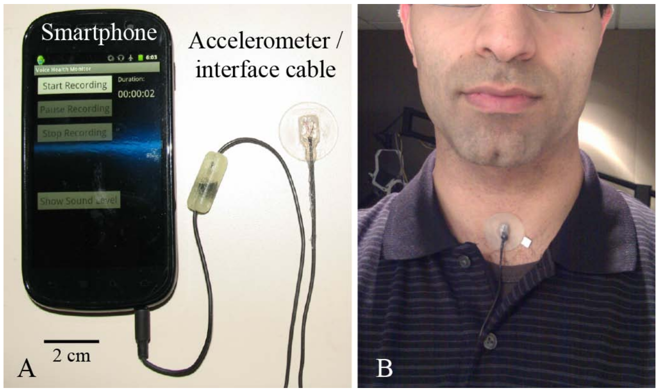


Fig. 2. Voice Health Monitor: (A) Smartphone, accelerometer sensor, and interface cable with circuit encased in epoxy; (B) the wired accelerometer mounted on a silicone pad affixed to the neck midway between the Adam's apple and V-shaped notch of the collarbone.

3) /pae pae pae pae pae/ at comfortable, soft, and loud levels

The sustained cardinal vowels /a/, /i/, and /u/ aid in defining the subject's vowel space and allowing for stability/perturbation analysis on amplitude and frequency. The Rainbow Passage is a standard paragraph in clinical voice assessment. Finally the string of /pae/ syllables is designed to elicit stable periods of intraoral pressure that equilibrate with lung pressure to provide indirect estimates of subglottal driving pressure.

In addition, we ask vocally normal subjects to sustain vowels /e/ and /o/ to provide a more complete vowel space. Also, the five vowels and /pae/ syllables are produced using three voice qualities that simulate effects of hyperfunctional voice use—breathy, pressed, and rough voice qualities.

### *B. In-field calibration and monitoring*

In the field, the device provides a user-friendly interface for voice health monitoring, daily sensor calibration, and periodic alert capabilities. Each morning, the subject is taken through a daily calibration sequence that seeks to match accelerometer signal levels to acoustic signal levels [9]. A simple linear regression between the energy of the neck acceleration and microphone signals from soft-to-loud vowel gestures are used to obtain estimates of the relationship between the two signals. A custom metal standoff maintains a fixed distance of 15 cm between the subject's lips and a handheld audio recorder (H1 Handy Recorder, Zoom Corporation, Tokyo, Japan) while the subject produces sustained /a/ vowels and reads the Rainbow Passage. Such daily calibrations can enhance the accuracy of accelerometer-based estimates of vocal dose measures and provide verification that the system is operating properly.

The subject affixes the accelerometer assembly to his or her neck using hypoallergenic double-sided tape (Model 2181, 3M, Maplewood, MN). The tape's circular shape and small tab allow for easy placement and removal of the accelerometer on the neck skin a few centimeters above the collarbone. The tape is strong enough to hold the silicone pad in place during a full day for the typical user.

## III. Subject recruitment and analysis approaches

The current phase of the project uses the new voice monitor to collect long-term ambulatory accelerometer data on groups of patients with hyperfunctional voice disorders (before and after treatment) and matched controls. Two groups of subjects with voice disorders are being enrolled: 1) patients with adducted hyperfunction (e.g., vocal fold nodules or polyps) and 2) patients with non-adducted hyperfunction (muscle tension dysphonia). Diagnoses are based on a complete team evaluation (laryngologist and speech-language pathologist) at the MGH Voice Center that includes 1) collection of a complete case history, 2) endoscopic imaging of the larynx, 3) Voice-Related Quality of Life (V-RQOL) questionnaire [10], 4) Consensus Auditory Perception of Voice Quality (CAPE-V) assessment [11], and 5) aerodynamic and acoustic assessment of vocal function. All of the subjects recruited are typically engaged in occupations that are considered to be at a higher-than-normal risk for developing a voice disorder [12]. Example occupations include singer, counselor/social worker, teacher, lawyer, clergy, telemarketer, salesperson, and health-care worker.

Two associated matched-control groups are obtained for each of the two patient groups. Each patient helps identify a work colleague of the same gender who has a normal voice and approximately the same age (within 5 years). The normal vocal status of all control subjects are then verified via interview and a laryngeal stroboscopic examination; otherwise, the subjects are screened out of the study. Each control subject is monitored for one full 7-day week.

Patients are monitored for one week at a time at particular time points during their treatment. Patients with adducted hyperfunction are monitored depending on the clinical management of their vocal fold lesions. At the MGH Voice Center, approximately 50 % of these patients undergo surgery followed by voice therapy, whereas the remaining 50 % receive only voice therapy. Thus approximately half of patients with adducted hyperfunction are expected to be monitored for three nonconsecutive weeks (pre-surgery, post-surgery, and post-

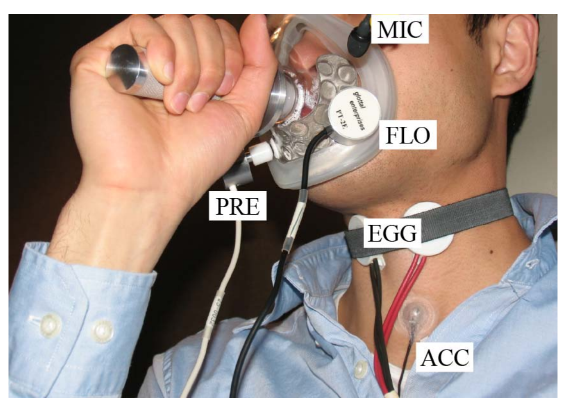


Fig. 3. In-laboratory calibration setup. Synchronized recordings are made of signals using an acoustic microphone (MIC), electroglottography electrodes (EGG), accelerometer sensor (ACC), radiated high-bandwidth airflow (FLO), and intraoral pressure (PRE).

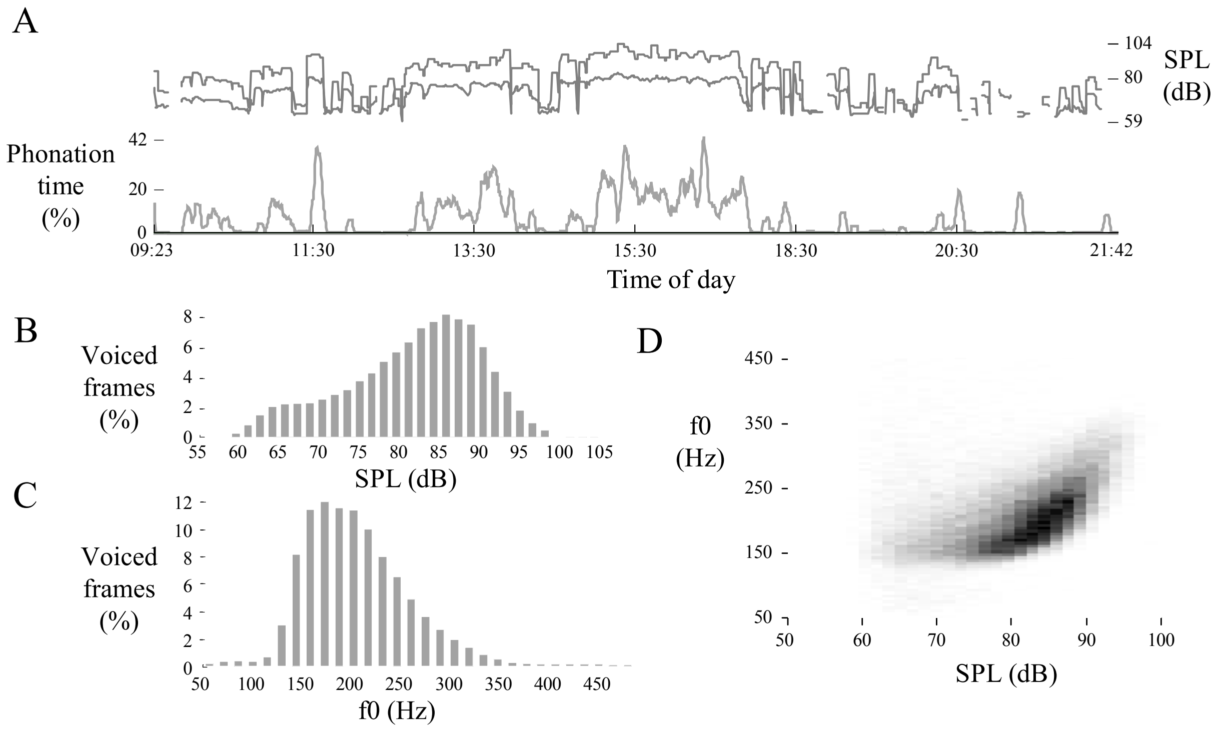


Fig. 4. Voice use profile of the subject in Figure 1B with adducted hyperfunction during her third day of monitoring (24-hour time format): (A) Five-minute moving average of percent phonation time, average sound pressure level (SPL; lower line), and maximum SPL (upper line) for voiced frames; (B) histogram of fundamental frequency (f0); (C) histogram of SPL; and (D) phonation density showing the relative occurrence of particular combinations of SPL (horizontal axis) and f0 (vertical axis).

therapy) and the other half are expected to be monitored for two nonconsecutive weeks (pre-therapy and post-therapy). Patients with non-adducted hyperfunction do not have lesions and are thus always monitored for two nonconsecutive weeks (pre-therapy and post-therapy).

### A. Vocal dose measures

Three vocal dose measures [13]—phonation time, cycle dose, and distance dose—are highlighted in the current study to quantify accumulated daily voice use for each subject. Phonation (voicing) time reflects the duration of frames containing vocal fold vibration and is computed in terms of time (hours:minutes:seconds) and percentage of total time. The cycle dose is an estimate of the number of vocal fold oscillations during a given period of time. Finally, the distance dose estimates the total distance traveled by the vocal folds, combining cycle dose with estimates of vibratory amplitude based on acoustic intensity.

Figure 4 displays an example of a voice use profile of the subject with adducted hyperfunction in Figure 1B. Such visualizations may ultimately enable clinicians to make informed decisions regarding the management or prevention of pathologies that could arise due to certain patterns of voice use in terms of sound pressure level and fundamental frequency. Note that the subject uses her voice 20–40 % of the time during the afternoon hours, which are values exceeding even those of teachers who have significant vocal demands during the work day [14].

### B. Mathematical modeling of accelerometer signal

Airflow-based measures of voice production greatly enhance the potential of ambulatory monitoring to detect some of the features that have previously been shown to differentiate among hyperfunctionally related disorders and normal modes of vocal function [3], [4]. Robust automatic detection of these parameters, however, has yet to be achieved in an ambulatory voice monitoring system. To address this goal, a model-based inverse filtering scheme, known as subglottal impedance-based inverse filtering (IBIF) [8], [15], is applied to produce glottal airflow measures from the neck surface acceleration signal recorded by the Nexus S smartphone.

Skin parameters for the subglottal IBIF model only need to be obtained once, often most readily achieved in the sustained vowel context. The airflow measures derived from the model are compared with estimates derived from the oral airflow signal recorded by the pneumotachograph mask system. Inverse filtering [16] of the oral airflow yields reference estimates of the glottal airflow waveform, making use of the synchronous electroglottographic signal to aid in identifying glottal closure instants. Processing the accelerometer signal can provide useful estimates of several glottal airflow properties, such as spectral slope (H1–H2), maximum flow declination rate (MFDR), amplitude of the modulated flow component (AC Flow), and open quotient (OQ). Details on these measures and their clinical relevance are available in the literature [3], [17], [18].

TABLE I

VOCAL DOSE MEASURES OF SUBJECTS WHOSE VOCAL FOLDS WERE SHOWN IN FIGURE 1. VALUES REPORTED ARE AVERAGES (STANDARD DEVIATIONS) OF DAILY STATISTICS OVER SEVEN DAYS FOR EACH SUBJECT.

| **Measure** | **Normal** | | **Adducted hyperfunction** | | **Non-adducted hyperfunction** | |
|---|---|---|---|---|---|---|
| Total time (hh:mm:ss) | 16:26:43 | (02:09:30) | 12:16:27 | (02:04:44) | 10:49:16 | (03:33:12) |
| Phonation time (hh:mm:ss) | 01:03:58 | (00:32:15) | 01:16:29 | (00:34:38) | 00:38:12 | (00:16:28) |
| Percent phonation (%) | 6.54 | (3.44) | 10.31 | (3.78) | 5.72 | (1.49) |
| F0 average (Hz) | 230.3 | (24.0) | 186.3 | (12.8) | 244.1 | (45.7) |
| Sound level (dB SPL) | 76.1 | (3.4) | 78.5 | (1.3) | 72.7 | (1.8) |
| Cycle dose (cycles) | 891,620 | (479,157) | 848,519 | (369,733) | 590,763 | (291,074) |
| Distance dose (m) | 2985.1 | (1664.2) | 3376.8 | (1477.4) | 1361.4 | (646.8) |

## IV. Initial results

Data collection is under way with the goal of enrolling a total of 400 subjects—200 subjects with vocal hyperfunction and 200 matched controls. The purpose of this section is to present initial results from the current data set to demonstrate a general framework for analyzing the long-term monitoring data that can be used on a larger sample to achieve statistically significant discrimination.

### *A. Vocal dose summary statistics*

Table I displays vocal dose measures for the adult female subjects whose vocal folds were shown in Figure 1. The vocally normal subject appears to exhibit lower percent phonation time and distance dose measures compared to corresponding measures from the adducted hyperfunction patient (with a vocal fold polyp), even though the normal subject's cycle dose is higher. The non-adducted hyperfunction subject appears to exhibit lower percent phonation time and vocal dose measures than corresponding measures from both control and adducted hyperfunction subjects. As additional subjects are enrolled in each subject group, these average vocal dose measures may reveal further discriminatory power, especially when coupled with phonation profiles (as in Figure 4) that display the time variation and distribution of fundamental frequency, sound pressure level, and phonation time.

### *B. Model-based measures of airflow*

Figure 5 illustrates signals and selected measures during Rainbow Passage reading derived from the accelerometer signal using subglottal IBIF, side by side with measures estimated from the inverse filtered oral airflow. Shown are data from an adult female with no voice disorder and her matched subject diagnosed with adducted hyperfunction. It is noted from the trajectories that the IBIF-based signals appear to be more stable and less prone to inverse filtering artifacts than their oral airflow–based counterparts. The accelerometer signal has a simplified dynamic behavior relative to the oral airflow, which includes time-varying resonances and constitutes a more challenging inverse filtering situation. As a result, the measures derived from the voice health monitor during running speech appear to be more reliable than those from more standard acoustic methods.

The MFDR and AC Flow measures of the hyperfunctional subject are elevated compared to her matched control derived from both accelerometer and oral airflow signals. MFDR and AC Flow appear to discriminate the subjects. The other measures either do not show differences as clearly or they have variations that are too large to distinguish the subjects, which is in agreement with previous studies of hyperfunctional voices that were based on sustained vowels [3].

## V. Summary and future directions

This paper provided an overview of a voice health monitor that uses a wearable sensor to record and characterize the phonation-related acceleration signal of the neck skin just above the collarbone. In-laboratory and in-field calibration protocols provide multisensor reference points from which waveform features relevant to clinical voice assessment are derived. Two signal processing approaches yield standard vocal dose measures and novel model-based estimates of glottal airflow characteristics. Initial results from these approaches provide an analysis framework with which the most common types of hyperfunctionally related voice disorders can be tracked over the course of days as speakers go about their typical routines. An understanding of daily behavior is essential to improving the diagnosis and treatment of hyperfunctional voice disorders.

Alternative analysis approaches could include spectral- and cepstral-based measures that have been shown to correlate with voice quality and level of dysphonia [19] and that would have to be translated to apply to the accelerometer signal. We will continue to refine algorithms for dynamically extracting airflow-based parameters in running speech from the accelerometer signal to optimize the detection of hyperfunctional vocal behaviors. In addition to such model-based approaches, we plan to explore the application of symbolic analysis techniques that have been previously used in the context of long-term cardiovascular monitoring for risk stratification of patients with cardiac disorders [20], [21]. With large sampling of long-term subject recordings, advanced machine learning techniques such as support vector machines could also be applied to help differentiate among hyperfunctional and normal modes of daily vocal behavior. Preliminary work on the classification of normal voice modes and simulated vocal hyperfunction by a trained speaker shows promise for delineating/targeting abusive vocal patterns during the course of voice therapy [22].

Wearable voice monitoring systems have the potential to provide more reliable and objective measures of voice use that can enhance the treatment of hyperfunctional voice disorders

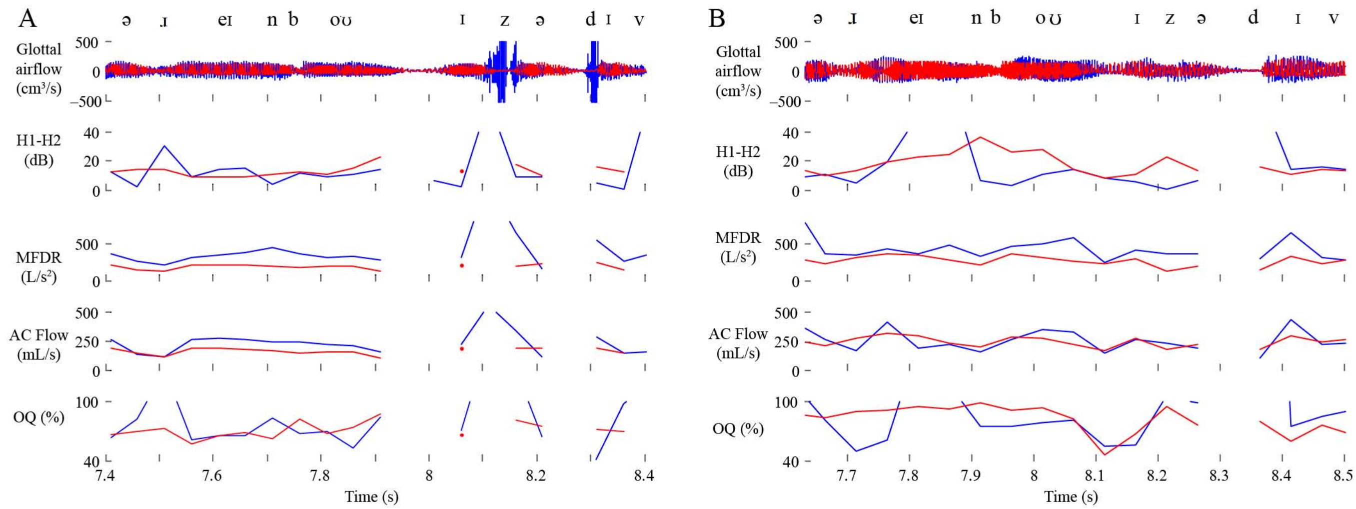


Fig. 5. Time-varying estimation of features of merit derived from the inverse-filtered airflow signal (blue) and the IBIF model–based glottal airflow signal (red). Trajectories are shown for (A) an adult female with no vocal pathology and (B) her matched patient diagnosed with adducted hyperfunction and bilateral vocal fold nodules. The phonetic transcription is shown above each waveform.

by unobtrusively delineating detrimental vocal behaviors that can be targeted for voice therapy treatment and, ultimately, provide real-time feedback to a speaker to facilitate the adoption of healthier vocal behaviors into every day use.

## Acknowledgment

This work was supported by the NIH National Institute on Deafness and Other Communication Disorders under grants R33 DC011588 and T32 DC00038, Universidad Técnica Federico Santa María, the Chilean CONICYT under Grant FONDECYT 11110147, and the MIT International Science and Technology Initiatives (MISTI) MIT-Chile Seed Fund under grant 2745333. The authors would like to thank Rob Petit for programming the Android application.

## References

[1] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith, "Voice disorders in the general population: Prevalence, risk factors, and occupational impact," *Laryngoscope*, vol. 115, no. 11, pp. 1988–1995, 2005.

[2] National Institute on Deafness and Other Communication Disorders (NIDCD), "2012-2016 strategic plan," *National Institutes of Health, U.S. Department of Health and Human Services*, no. 6, pp. 1–63, 2012.

[3] R. E. Hillman, E. B. Holmberg, J. S. Perkell, M. Walsh, and C. Vaughan, "Objective assessment of vocal hyperfunction: An experimental framework and initial results," *J. Speech Hear. Res.*, vol. 32, no. 2, pp. 373–392, 1989.

[4] ——, "Phonatory function associated with hyperfunctionally related vocal fold lesions," *J. Voice*, vol. 4, no. 1, pp. 52–63, 1990.

[5] H. A. Cheyne, H. M. Hanson, R. P. Genereux, K. N. Stevens, and R. E. Hillman, "Development and testing of a portable vocal accumulator," *J. Speech. Lang. Hear. Res.*, vol. 46, no. 6, pp. 1457–1467, 2003.

[6] R. E. Hillman, J. T. Heaton, A. Masaki, S. M. Zeitels, and H. A. Cheyne, "Ambulatory monitoring of disordered voices," *Ann. Otol. Rhinol. Laryngol.*, vol. 115, no. 11, pp. 795–801, 2006.

[7] D. D. Mehta, M. Zañartu, S. W. Feng, H. A. Cheyne II, and R. E. Hillman, "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3090–3096, 2012.

[8] M. Zañartu, J. C. Ho, D. D. Mehta, R. E. Hillman, and G. R. Wodicka, "Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration," *IEEE Trans. Audio Speech Lang. Processing*, accepted, 2012.

[9] J. G. Švec, I. R. Titze, and P. S. Popolo, "Estimation of sound pressure levels of voiced speech from skin vibration of the neck," *J. Acoust. Soc. Am.*, vol. 117, no. 3, pp. 1386–1394, 2005.

[10] N. D. Hogikyan and G. Sethuraman, "Validation of an instrument to measure voice-related quality of life (V-RQOL)," *J. Voice*, vol. 13, no. 4, pp. 557–569, 1999.

[11] G. B. Kempster, B. R. Gerratt, K. Verdolini Abbott, J. Barkmeier-Kraemer, and R. E. Hillman, "Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol," *Am. J. Speech Lang. Pathol.*, vol. 18, no. 2, pp. 124–132, 2009.

[12] K. Verdolini and L. O. Ramig, "Review: Occupational risks for voice problems," *Log. Phon. Voc.*, vol. 26, no. 1, pp. 37–46, 2001.

[13] I. R. Titze, J. G. Švec, and P. S. Popolo, "Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues," *J. Speech. Lang. Hear. Res.*, vol. 46, no. 4, pp. 919–932, 2003.

[14] I. R. Titze, E. J. Hunter, and J. G. Švec, "Voicing and silence periods in daily and weekly vocalizations of teachers," *J. Acoust. Soc. Am.*, vol. 121, no. 1, pp. 469–478, 2007.

[15] M. Zañartu, "Acoustic coupling in phonation and its effect on inverse filtering of oral airflow and neck surface acceleration," Ph.D. dissertation, Purdue University, West Lafayette, IN, 2010.

[16] P. Alku, C. Magi, S. Yrttiaho, T. Backstrom, and B. Story, "Closed phase covariance analysis based on constrained linear prediction for glottal inverse filtering," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3289–3305, 2009.

[17] I. Titze, *Principles of Voice Production*. Iowa City, IA: National Center for Voice and Speech, 2000.

[18] D. D. Mehta and R. E. Hillman, "Voice assessment: Updates on perceptual, acoustic, aerodynamic, and endoscopic imaging methods," *Curr. Opin. Otolaryngol. Head Neck Surg.*, vol. 16, no. 3, pp. 211–215, 2008.

[19] S. N. Awan, N. Roy, M. E. Jetté, G. S. Meltzner, and R. E. Hillman, "Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V," *Clin. Linguist. Phon.*, vol. 24, no. 9, pp. 742–758, 2010.

[20] Z. Syed, J. Guttag, and C. Stultz, "Clustering and symbolic analysis of cardiovascular signals: Discovery and visualization of medically relevant patterns in long-term data using limited prior knowledge," *EURASIP J. Adv. Sig. Process.*, pp. 1–16, 2007.

[21] Z. Syed, P. Sung, B. M. Scirica, D. A. Morrow, C. M. Stultz, and J. V. Guttag, "Spectral energy of ECG morphologic differences to predict death," *Cardiovasc. Eng.*, vol. 9, no. 1, pp. 18–26, 2009.

[22] M. Ghassemi, E. Shih, D. Mehta, S. Feng, J. Van Stan, R. Hillman, and J. Guttag, "Detecting voice modes for vocal hyperfunction prevention," *Proceedings of the 7th Annual Workshop for Women in Machine Learning*, 2012.