Proceedings of the 10th International Conference on
Advances in Quantitative Laryngology, Voice and Speech Research
June 3-4, 2013, Cincinnati, Ohio, USA

FUTURE DIRECTIONS IN THE DEVELOPMENT OF AMBULATORY MONITORING FOR CLINICAL VOICE ASSESSMENT

Robert E. Hillman[1], Matías Zañartu[2], Marzyeh Ghassemi[3], Daryush D. Mehta[1], Jarrad H. Van Stan[1], Harold A. Cheyne II[4], John V. Guttag[3]

[1] Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA
[2] Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
[3] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA
[4] Bioacoustics Research Program, Cornell University, Ithaca, NY, USA

***Keywords*: Voice; Voice Disorders; Vocal Hyperfunction; Ambulatory Monitoring; Accelerometer**

INTRODUCTION

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* [1]. Such behaviorally based disorders can be difficult to accurately assess in the clinical setting and could potentially be much better characterized by long-term ambulatory monitoring of vocal function as individuals engage in their typical daily activities. Devices that use a neck-placed miniature accelerometer (ACC) as a phonation sensor have shown the good potential for unobtrusive long-term monitoring of vocal function. However, the adoption of this technology into clinical practice has been quite limited due mainly to the lack of statistically robust studies that demonstrate the true diagnostic capabilities of ACC-based measures.

This report describes the status and preliminary results for an ongoing project that is designed to develop ACC-based ambulatory monitoring of vocal function into a valid, reliable, and cost-effective clinical tool that can be used to accurately identify and differentiate patterns of voice use that are associated with hyperfunctional voice disorders. Achieving this goal should improve clinical assessment and treatment of these commonly-occurring types of voice disorders.

METHODS

Human Data

Data collection is aimed at obtaining a large, statistically robust sample of patients with hyperfunctional voice disorders (before and after treatment) and matched controls. This involves long-term ambulatory monitoring of four groups of 100 subjects each (total of 400 subjects): 1) patients with vocal fold nodules or polyps (adducted hyperfunction), 2) patients with muscle tension dysphonia in the absence of vocal fold lesions (non-adducted hyperfunction), and 3) two associated matched-control groups with normal status confirmed by endoscopic videstroboscopy. In order to obtain matched controls for each of the two disordered groups, each voice patient that is enrolled in the study is asked to identify a colleague/co-worker who has a normal voice, the same gender, same approximate age (within 5 years), similar occupation, and is willing to wear the voice monitor for one week.

All subjects are monitored all day for a minimum of 7 consecutive days. Patients with voice disorders are monitored for 1 week prior to treatment and then for 1 additional week following any treatments that they receive (surgery and/or voice therapy).

Instrumentation and Measurement

Each data collection system is comprised of a Nexus S smartphone (running the Android operating system and a custom-designed software application) and a neck-placed ACC that is connected to the phone by a wiring assembly that includes a miniaturized interface circuit and a jack that plugs into the regular input socket of the phone [2].

The data collection system is calibrated relative to acoustic (microphone) and aerodynamic (Rothenberg mask) parameters for each subject prior to obtaining ambulatory recordings. The system records the raw ACC signal, and its application provides capabilities for recording starting/stopping, in-field acoustic calibration and system checks, and the collection of additional subject response data (multiple choice questions and linear analog scales) concerning daily activity and vocal status queries presented on the screen of the phone.

Analysis

The ambulatory ACC data are being subjected to three types of analysis approaches in an effort to identify the best set of measures for differentiating among hyperfunctional and normal patterns of vocal behavior:

1. Previously-developed ambulatory measures of vocal functions that are extracted directly from the ACC signal and include phonation time, fundamental frequency (f0), sound pressure level (SPL), and vocal doses (including cycle and distance doses).

2. Measures based on estimates of the modulated (AC) glottal airflow that are extracted from the ACC signal using a new vocal system [3]. Estimated measures include

air flow amplitude, maximum flow declination rate (MFDR), various cycle quotients (e.g., speed and open quotients), and spectral measures (e.g., H1–H2).

3. Use of machine learning techniques to identify features in long-term ACC recordings that are indicative/predictive of vocal status (normal vs. hyperfunction/pathological).

Results and Discussion

Data collection is well underway (> 50 subjects recorded), and the direct and model-based AC flow measures have been refined and automated. The system for automatically extracting previously-developed measures of vocal function from the ACC signal has been finalized, and an example of its graphical output is shown in Fig. 1. These results are used as the initial step in checking the quality of each daily recording and generating data for later group-based statistical analyses.

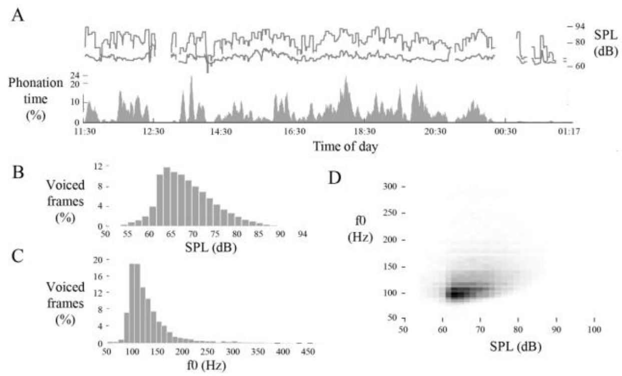

Fig. 1: Example of graphical output for previously-developed ambulatory monitoring measures.

The initial version of the system for extracting AC flow–based measures from the ACC signal during running speech (using our new approach of subglottal impedance-based inverse filtering, IBIF) has been developed. Table 1 shows comparisons of selected measures for a female patient with vocal nodules (adducted hyperfunction) and her matched control derived from running speech using both subglottal IBIF of the ACC signal and traditional inverse filtering of the oral volume velocity airflow (OVV).

Table 1: Mean (standard deviation) of selected glottal parameters during running speech from ACC and OVV signals for a female with nodules and her matched control. Measurement units: H1–H2 (dB), MFDR (L/s^2), AC Flow (mL/s), OQ (%).

Measure	Vocal Nodules		Matched Control	
	ACC	OVV	ACC	OVV
H1–H2	12.2 (4.8)	16.1 (15.5)	12.2 (4.0)	11.5 (11.0)
MFDR	286.7 (94.8)	675 (1237.4)	186.2 (62.4)	370.4 (339.7)
AC Flow	260.9 (65.7)	408.1 (509.7)	149.0 (41.7)	235.7 (149.1)
OQ	84.2 (143.2)	145.0 (716.7)	69.2 (25.2)	75.7 (72.3)

The IBIF-based signals and measures appear more stable and less prone to inverse filtering artifacts than their oral airflow–based counterparts, as evidenced by smaller standard deviations.

The increased stability of ACC-based measures is potentially due to the more simplified dynamic behavior of the ACC signal relative to the oral airflow that includes time-varying resonances and presents a more challenging problem for inverse filtering. The MFDR and AC Flow values for the nodules patient are much higher than those for the control subject, which agree with previous observations of hyperfunctional voices that were based on inverse filtering oral airflow [1].

A preliminary test of machine learning techniques was performed using data from 6 patients with vocal nodules (adducted hyperfunction) and their 6 matched controls. Initial results using both adjusted logistic regression and support vector machine approaches are consistent with previous observations that long-term measurement averages (e.g., phonation time, SPL, distance dose) do not consistently differentiate between normal and hyperfunctional/pathological conditions. Evidence of much better (statistically significant) differentiation is found when examining features that are derived to capture the incidence of more extreme vocal behaviors.

Conclusion

Excellent progress is being made in developing ACC-based ambulatory monitoring of vocal function into a valid, reliable, and cost-effective clinical tool that can be used to accurately identify and differentiate patterns of voice use that are associated with hyperfunctional voice disorders. This work will continue to focus on further improvement of measures and the collection of larger data sets to increase the statistical power of test results.

Acknowledgments

Research funded by NIH grant 5R33DC011588-03, CONICYT grant FONDECYT 11110147, and MIT-Chile MISTI grant 2745333.

References

[1] 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;32:373–392.

[2] 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;59:3090–3096.

[3] 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;accepted for publication.