INCORPORATING REAL-TIME BIOFEEDBACK CAPABILITIES INTO A VOICE HEALTH MONITOR

Andrés F. Llico\textsuperscript{1}, Matías Zañartu\textsuperscript{1*}, Daryush D. Mehta\textsuperscript{2}, Jarrad H. Van Stan\textsuperscript{2}, Harold A. Cheyne II\textsuperscript{3}, Agustín J. González\textsuperscript{1}, Marzyeh Ghassemi\textsuperscript{4}, George R. Wodicka\textsuperscript{5}, John V. Guttag\textsuperscript{4}, and Robert E. Hillman\textsuperscript{2}

\textsuperscript{1}Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
\textsuperscript{2}Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA
\textsuperscript{3}Bioacoustics Research Program, Cornell University, Ithaca, NY, USA
\textsuperscript{4}Computer Science and Artificial Intelligence Lab., Massachusetts Institute of Technology, Cambridge, MA, USA
\textsuperscript{5}Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

* Corresponding author. Email: matias.zanartu@usm.cl

I. INTRODUCTION

It is believed that abusive or faulty patterns of vocal behavior lead to phonotraumatic lesions on the vocal folds such as nodules and polyps or functional dysphonia. Recent research suggests that these behaviors, often referred to as vocal hyperfunction, might be better characterized through long-term ambulatory voice monitoring during a patient’s daily activities than with traditional clinical assessment methods [1]. Such an ambulatory monitoring system that employed a neck surface accelerometer (ACC) was developed by Cheyne and colleagues and was shown to provide a number of advantages compared to microphone-based systems [2]. This device, the Ambulatory Phonation Monitor (APM), is commercially-available for research and clinical use (Model 3200, KayPENTAX) and records only vocal parameters, i.e., fundamental frequency (F0) and sound pressure level (SPL) every 50 ms for an entire day. The APM also provides biofeedback capabilities (via a separate pager vibrator) based on upper or lower thresholds set for F0 and/or SPL. Ambulatory biofeedback using the APM has been shown in early case studies to have the potential to facilitate vocal behavioral changes targeted in voice therapy [1]. An enhanced ambulatory system under development, referred to as the Voice Health Monitor (VHM), incorporates an ACC and uses a smartphone platform [3], as shown in Fig. 1a and 1b. The VHM has 80 dB dynamic range, overcomes the storage limitations of the APM (thereby allowing for recording raw acceleration data for at least 7 days), and provides an enhanced processing platform to run complex algorithms [3]. Prior to this study, the VHM operated only as a data acquisition system, without biofeedback capabilities. This study aims to expand the VHM capability by incorporating real-time biofeedback capabilities and comparing its performance with that of the APM.

II. METHODS

Given the processing capabilities of the smartphone platform, numerous biofeedback approaches can be implemented in the VHM. In this study, we focus on the real-time estimation of F0 and SPL in order to mimic and contrast the current real-time biofeedback capabilities of the APM. This comparison is performed using a repeatable excitation signal and biofeedback triggering setup. As shown in Fig. 1c, the same light-weight accelerometer provides the input stimulus for both systems. The accelerometer was mounted on a bioacoustic transducer tester (BATT) [4] that was set to have a flat, band-limited response between 70 Hz and 2 kHz. The BATT was excited with an ambulatory recording previously captured with the VHM from an adult male subject with normal voice (a teacher during a 90-minute lecture), thus providing a signal comparable to that initially obtained with the VHM. Both systems were calibrated with the same subject-specific parameters, initially optimized for the APM. Only voiced frames that exhibited SPL ≥ 62 dB and 50 Hz ≤ F0 ≤ 500 Hz were considered. Biofeedback was triggered at SPL ≥ 95 dB with 300ms onset.

![Fig. 1. Mobile Voice Health Monitoring System: (a) Smartphone, in-line interface circuit, and accelerometer input, (b) Subject wearing the VHM system, and (c) BATT platform for APM and VHM comparisons.](image-url)
### III. RESULTS AND DISCUSSION

The summary statistics for each device are shown in Table I. Although the same conditions and calibration were provided for each, the measured phonation time, percentage of compliance, and biofeedback time differ slightly. Average differences of F0 and SPL estimates were less than 1 Hz and 1 dB, respectively. Although the average behaviors were similar, further differences were observed from the parameter histograms shown in Fig. 2. The slightly greater APM values around the average F0 in Fig. 2a are consistent with the APM labeling more frames as voiced than did the VHM. The SPL histograms in Fig. 2b show that most differences were in around both the left tails and modes, hence producing equivalent SPL averages. These findings indicate that the APM tends to label more lower-energy frames as voiced and shifts the center of the distribution, thus explaining its increased accumulated phonation time, lower compliance time and higher biofeedback time. Although the differences under the testing conditions are less than 1%, the results for the VHM better match those reported in the literature [1-3,5].

<table>
<thead>
<tr>
<th>Device</th>
<th>Total Time (hh:mm:ss)</th>
<th>Phonation Time (hh:mm:ss)</th>
<th>Average F0 (Hz)</th>
<th>Average SPL (dB)</th>
<th>% Compliance (SPL &lt;= 95 dB)</th>
<th>Biofeedback Time (hh:mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APM</td>
<td>02:08:47</td>
<td>00:38:34 (29.96%)</td>
<td>144.9</td>
<td>82.4</td>
<td>93.6</td>
<td>00:00:45 (1.98 %)</td>
</tr>
<tr>
<td>VHM</td>
<td>02:08:52</td>
<td>00:32:39 (25.34%)</td>
<td>145.7</td>
<td>81.7</td>
<td>95.7</td>
<td>00:00:20 (1.04 %)</td>
</tr>
</tbody>
</table>

Fig. 2. Histograms for APM (blue) and VHM (red): (a) F0 and (b) SPL.

Due to differences in memory allocations to amplitude values, the VHM has about 40 dB more dynamic range than the APM. This may explain the differences in SPL and would indicate that the APM is less reliable when representing a wide range of SPL (which is in accordance with our clinical observations). Therefore, the added real-time biofeedback capabilities in the VHM are considered more precise and better suited for singers or other patients that have large voice range profiles. Subsequent investigations with the VHM in the context of an enhanced real-time biofeedback include the estimation of aerodynamic parameters using impedance-based inverse filtering [6], cepstral peak prominence, relative fundamental frequency, and wireless connectivity with a server in the clinic. The ability to better facilitate vocal behavioral changes with these new real-time features remains to be tested.

### ACKNOWLEDGMENT

This work was supported by NIH-NIDCD grant R33 DC011588, CONICYT grant FONDECYT 11110147, and MIT MISTI grant MIT-Chile 2745333.

### REFERENCES