Lightweight, On-Body, Wireless System for Ambulatory Voice and Ambient Noise Monitoring

Patrick C. Chwalek, Daryush D. Mehta, *Member, IEEE*, Brendon Welsh, Catherine Wooten, Kate Byrd, Edward Froehlich, David Maurer, Joseph Lacirignola, Thomas F. Quatieri, *Fellow, IEEE*, Laura J. Brattain

***Abstract*— In this paper, we present a lightweight, on-body, wireless system designed for monitoring real-world voice characteristics and behavior. The system has the potential to provide important assessments of voice and speech disorders, psychological and emotional state, and the impact of environmental sound levels. The system's transmitter is positioned on the neck and synchronously streams dual-channel sensor data from an on-board MEMS microphone and a high-bandwidth accelerometer, which acts as a noise-robust and confidential contact microphone. These data are recorded to a receiver that can store the data locally and stream a real-time feed to a computer. We also report on the design considerations of this novel system and discuss progress leading up to the latest iteration, especially of the transmitter components on a flexible circuit.**

I. Introduction

In this paper, we discuss a wireless voice monitor system that we developed for recording synchronized acoustic (MEMS microphone; MIC) and non-acoustic (accelerometer; ACC) data. Our system offers a tether-less method of capturing voice-related features that are important for speaker identification, noise reduction, and, most notably, for exploiting non-acoustic vocal signatures in real-world environments to provide long-duration monitoring and real-time biofeedback. Since naturalistic environments make it challenging to estimate many important voice characteristics in noisy conditions, recordings of neck-surface vibration have been the subject of ongoing investigation due to their robustness to acoustic environmental noise, low profile, and lack of speech intelligibility (alleviating confidentiality concerns) [1, 2]. However, MIC recordings continue to be desirable to capture the airborne acoustic signal that can be analyzed to quantify speech features (e.g., formants) and environmental characteristics. For clinical testing, it is also desirable to have a small form-factor device so that it minimizes discomfort and leads to increased compliance using the device. Furthermore, we discuss designing a system of this size with a wireless communication requirement on a flexible substrate. There are several improvements discussed that were made to this current system from the previously reported version that was used to evaluate the Lombard effect in a laboratory setting with increasing background noise level [3].

II. Evolution of Ambulatory Voice Monitoring

In the laboratory, acoustic and non-acoustic vocal sensors have been applied previously to robustly characterize speech in the presence of various types of background noise by fusing features from multiple sensors, including a bone conduction MIC, radar sensor, and contact MICs [4]. In the field, ambulatory voice monitoring devices have typically incorporated either acoustic MICs, contact MICs, or high-bandwidth ACC sensors to capture vocal characteristics and behavior [5]. The primary application of such devices was the clinical assessment of voice use in the daily lives of individuals working in occupations associated with higher incidences of voice-related complaints. These technologies have allowed for the computation of parameters that have been associated with heavy voice use (increased talk time, inappropriate pitch and loudness, etc.) [6].

Recent work has developed a smartphone-based platform for ambulatory voice monitoring that records the raw signal from a wired high-bandwidth ACC, allowing for the exploration of novel parameters related to the acoustics and aerodynamics of vocal function [7, 8]. In addition, real-time computation of voice features on the smartphone enables the study of more sophisticated biofeedback scheduling known to be critical for behavior modification [9].

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited. This material is based upon work supported by the Assistant Secretary of Defense for Research and Engineering under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Assistant Secretary of Defense for Research and Engineering.

P. C. Chwalek, D. D. Mehta, B. Welsh, C. Wooten, K. M. Byrd, E. Froehlich, D. Maurer, J. Lacirignola, T. F. Quatieri, and L. J. Brattain are with MIT Lincoln Laboratory, Lexington, MA 02421 USA (e-mail: [patrick.chwalek, daryush.mehta, brendon.welsh, catherine.wooten, kate.byrd, edward.froehlich, david.maurer, lacirignola, quatieri, brattainl]@ll.mit.edu). Corresponding author email: brattainl@ll.mit.edu.

D. D. Mehta is also with Massachusetts General Hospital, Harvard Medical School, and the MGH Institute of Health Professions, Boston, MA USA (e-mail: mehta.daryush@mgh.harvard.edu).

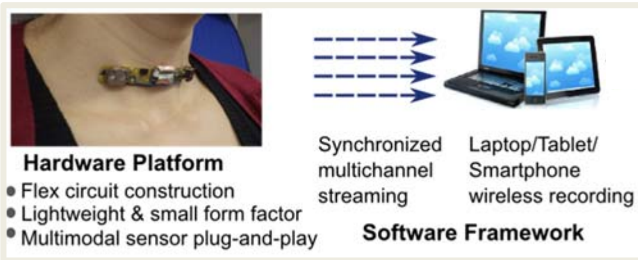

Figure 1. System framework of wireless voice monitor.

All previous devices, however, required a tethered connection which can be cumbersome during data collection. Users may snag cabling while walking or find wiring uncomfortable during long-term use, leading to user non-compliance and the potential for noisy recordings. Thus, our system's wireless design was a high priority, as well as the ability to program user-specific algorithms that can perform computations on real-time data streams captured by the receiver (see Fig. 1). Like the VoxLog device, we maintain the desirable dual-channel recording of both acoustic MIC and high-bandwidth ACC sensors for the noise-robust monitoring of voice use in naturalistic environments. An early prototype of our wireless system was initially reported as part of a comprehensive multimodal system for animal behavior monitoring [10]; the module continues to undergo significant improvements and optimization for human voice analysis [3], with the latest version being the subject of the current paper.

III. System Components

The system consists of two components: (1) a wearable, on-body transmitter that is worn around the neck below the thyroid prominence and (2) a wireless receiver that can either be carried on body, in a pocket, or plugged into a computer to visualize raw signal streams. A summary of the system specifications is shown in Table 1.

A. *Wireless Transmitter*

As in the previous system iteration [3], the wireless transmitter (Fig. 2) consists of several components soldered on a flexible circuit substrate. For sound level analysis, we use an omnidirectional MEMS MIC (SPA2410LR5H-B, Knowles Electronics, Itasca, IL) because of its small form factor and wide bandwidth that includes frequencies that range from 100 Hz to 15 kHz. For non-acoustic voice sensing, we use a single-axis, high-bandwidth ACC (BU-27135, Knowles Electronics). Our improved circuit includes multiple amplification stages before the analog sensor signals are digitized by our transceiver to increase the signal-to-noise ratio (SNR).

TABLE I. Wireless Voice Monitor Specifications for the Dual-Channel Streaming of Microphone (MIC) and Accelerometer (ACC) Sensors

Feature	Specification
Sample rate	44.1 kHz (per channel)
Resolution	16 bits
Bandwidth	ACC: 0–5 kHz, MIC: 100 Hz–15 kHz
MIC Sensitivity	−38dB ± 3dB @ 94dB SPL
ACC Sensitivity	−45 dB @ 1kHz
Power Consumption (active streaming)	56 mW (transmitter) 330 mW (receiver)
Weight (circuit)	5 g (transmitter), 14 g w/ 0.4 Ah battery 13.5 g (receiver), 47.5 g w/ 1 Ah battery
Size (circuit)	Transmitter: 68 × 14.5 × 5 mm Receiver: 68 × 14.5 × 5 mm
Wireless Protocol	Bluetooth 4.0

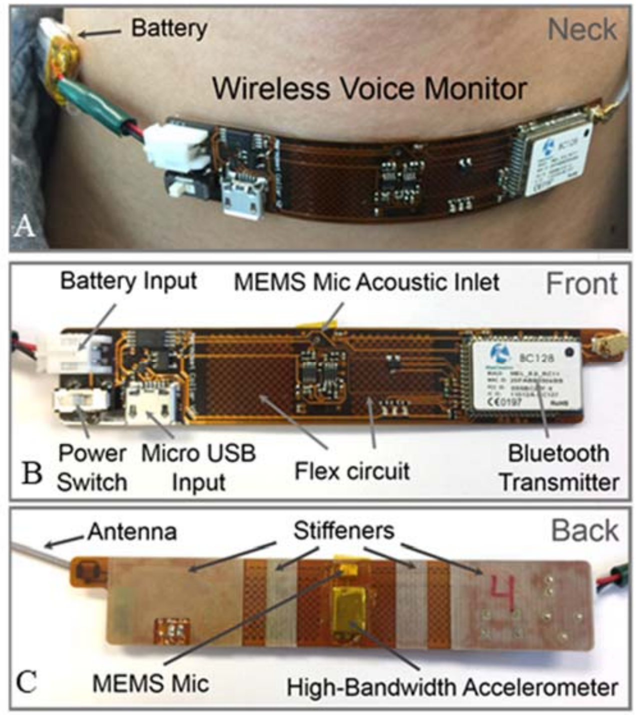

Figure 2. Wireless voice monitor, showing (A) transmitter circuitry on the anterior neck surface, (B) components on the front layer of the flexible circuit, and (C) components on the back layer.

The BC128 Bluetooth module (BlueCreation, Cambridge, UK) is the chosen transceiver due to its low power consumption, software configurability, 16-bit analog-to-digital converters, and ability to use an external antenna to minimize radio frequency (RF) noise inherent in a miniaturized printed circuit board (PCB). RF noise issues were prevalent in the previous system iteration [3] that included a BC127 module, which consisted of a built-in antenna that we found subjected the small circuit to a substantial amount of RF noise. The BC128 also offers several software features that can be toggled depending on the application (e.g., sending channel data to a Bluetooth-enabled phone, streaming to multiple receivers, etc.).

The ACC, MIC, and active circuit components are powered by a single-cell, rechargeable, lithium-ion polymer battery that can be charged through a micro-USB port on the circuit. The micro-USB input also allows for communication to the BC128 to modify firmware settings (e.g., channel gain) and troubleshoot. Additional features include electrostatic discharge protection, an on/off switch for the battery, status LEDs, and a power multiplexer integrated circuit that enables the BC128 to be fully functional when powered via USB and/or battery. The battery size is dependent on the application and the usual tradeoff of size and capacity. We chose a 400 mAh battery for the transmitter because of its small form factor and ability to run the system uninterrupted for at least 24 hours.

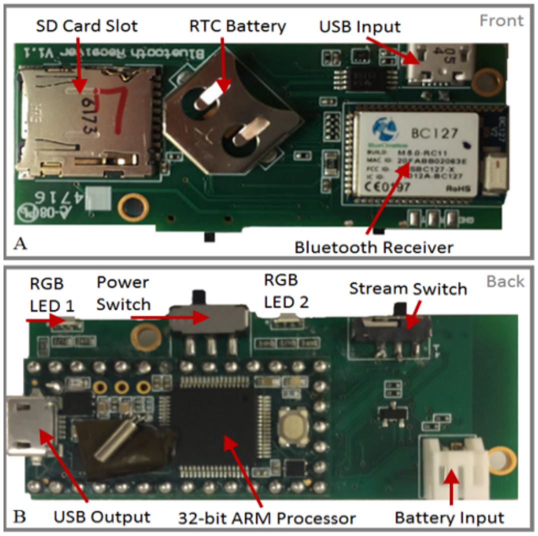

Figure 3. Wireless voice monitor's receiver. Components on the (A) front layer and (B) back layer of the circuit board.

B. Wireless Receiver

For data logging, we decided not to store the data on the neck-worn circuit since that would require additional space and power, making the neck-worn circuit bulkier, heavier, and less desirable for a user to wear throughout the day. We opted to design a separate receiver device, shown in Fig. 3, that can log the data sent over Bluetooth from the transmitter, as well as output the data streams in real-time to a personal computer (PC) via USB. We chose to build the device around a Teensy 3.2 microcontroller unit because of its small footprint, a 32-bit ARM processor with general-purpose input/output pins for future expansion, and active development community. For the Bluetooth receiver, we chose to use the BC127 module that has the same internal components as the BC128 but included the chip antenna in the same package. For the receiver, we did not have any RF-sensitive components so we found the BC127 ideal for a compact receiver design. In addition, the receiver has similar power regulation circuitry to the transmitter circuit, with additional circuitry that allows users to charge the receiver's larger battery at a faster rate. Red/green/blue LEDs on the receiver serve as system status indicators (recording status, Bluetooth pairing successful, etc.).

As mentioned above, the receiver module can be connected to a PC where the data can be streamed in real time from the transmitter module. The computer recognizes the USB input as a standard audio device, similar to a USB MIC. This allows the users to monitor data streams in real time. Regardless of whether real-time streaming to a PC is performed, the raw ACC and processed MIC data are saved on a Class 10 or higher micro SD card on the receiver module. We did not save raw MIC data due to the requirements of confidentiality during acoustic recording in a public environment. An illustration of the sensor waveforms streamed to PC and saved to SD card is shown in Fig. 4.

The receiver has an internal real-time clock that is powered by an external cell battery. When recording, the system saves the channel streams into five-minute data files, labeled with the starting time of the block. This scheme handles the case of a user being monitored in an environment that could interfere with the Bluetooth link; accurate timestamps would only be lost within that five-minute period. This scheme also solves issues of large data file preservation; if the system were to lose power inadvertently—e.g., due to a depleted battery—only up to the last five minutes of data would be lost due to incomplete closure of the last data file.

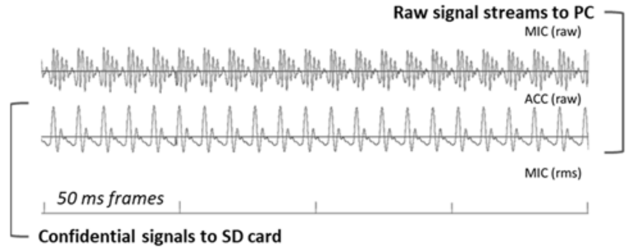

Figure 4. Sensor signals being transferred throughout the system. Laboratory data include raw ACC and MIC signals, whereas ambulatory recordings save data channels to preserve confidentiality (raw ACC and averaged MIC rms signals).

C. Design Considerations

Flexible circuit design has its challenges when compared to a rigid PCB, especially when building a relatively small form factor with sensitive analog circuitry that will be subjected to repetitive bending. Initially our ACC picked up noise artifacts from the MIC signal and from a preexisting chip antenna [3] so we opted for the BC128 with an external antenna that is mounted on the back of a user's neck. In addition, when designing the PCB, we employed various noise mitigation techniques to give us the best SNR. In addition to separating the ground planes between sensitive components, we added multiple operational amplifier stages placed next to the sensors to increase the SNR of the received signal at the BC128. The amplifier stages also allowed us to scale the dynamic ranges of both sensors separately, a feature lacking in the software-adjustable preamplifier in the BC128 module.

Flexible circuits are not designed to be often subjected to compressive and tensile stresses (i.e., multiple users taking the circuit off and on). To promote bending at certain regions while mitigating it at others where integrated circuits and passive components are mounted, we chose to vary our ground plane density. As can be seen from Fig. 2, we wanted to prevent bending where the sensors were co-located so we added a hatched copper pattern, versus a solid one underneath the sensors, on one side and omitted the copper plane on the

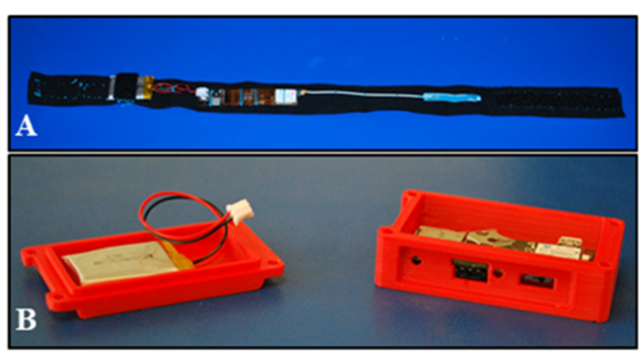

Figure 5. Packaged wireless voice monitor. (A) transmitter components sewn on fabric neck strap and (B) enclosed receiver.

other. In addition to localize bending, this method also provides a larger ground plane for electrical noise dissipation.

For packaging the transmitter (Fig. 5A), we chose a polyester material that would not irritate the user during long duration use. The polyester strap is commonly used for heart rate monitors strapped to a user's chest. To secure the device to the fabric and to further ruggedize it, we potted the circuit with Reprorubber® Thin Pour, a type of polysiloxane. During encasement, we masked off all ports and LEDs so that they remain accessible to the user. The battery can be swappable since different power capacities may be desired depending on the application (with the usual tradeoff of increased size).

For packaging the receiver (Fig. 5B), we required a case to protect sensitive circuitry from damage since it would be carried either on a belt clip or in a user's pocket/purse (and could potentially be dropped accidentally). Thus, we designed a rigid case that was 3D printed using Acrylonitrile Butadiene Styrene (ABS) plastic. The case fits the receiver board and a battery with enough capacity to last an entire day of data recording. Ports for USB inputs, switches, and LEDs were left open and accessible to the user.

IV. Conclusion

In this paper, we presented our latest development and implementation of an ambulatory voice monitor designed for the monitoring of real-world voice characteristics. The ability to stream synchronized sensor data from a MEMS MIC and an ACC on a flexible substrate allows us to correlate the two sensors. In addition, the acoustic MIC gives us complementary data when processing the raw non-acoustic signal. The ACC is more immune to acoustic noise and may reveal features not derivable from the MIC.

Future work could consist of adding a more capable processor to the receiver to enable more complex machine learning algorithms, Bluetooth communication with a smartphone, and additional sensors to the transmitter circuitry to increase the system's capabilities. Streaming more than two data streams would warrant shifting away from the off-the-shelf BC127/128 wireless modules and developing custom wireless transceiver circuitry and firmware.

Other vocal function measures available include ACC signal properties of periodicity, harmonic spectral tilt, low- to high-frequency spectral power ratio, and cepstral peak prominence [2]. These types of acoustic-based measures have known relationships with ACC signal properties [12] and can be explored along with aerodynamic measures of vocal function hypothesized to be particularly salient for the assessment of common voice disorders [2, 13, 14].

References

[1] M. Zañartu, J. C. Ho, S. S. Kraman, H. Pasterkamp, J. E. Huber, and G. R. Wodicka, "Air-borne and tissue-borne sensitivities of bioacoustic sensors used on the skin surface," *IEEE Transactions on Biomedical Engineering,* vol. 56, pp. 443-451, 2009.

[2] D. D. Mehta, J. H. Van Stan, M. Zañartu, M. Ghassemi, J. V. Guttag, V. M. Espinoza, *et al.*, "Using ambulatory voice monitoring to investigate common voice disorders: Research update," *Frontiers in Bioengineering and Biotechnology,* vol. 3, pp. 1-14, 2015.

[3] D. D. Mehta, P. C. Chwalek, T. F. Quatieri, and L. J. Brattain, "Wireless neck-surface accelerometer and microphone on flex circuit with application to noise-robust monitoring of Lombard speech," *Proceedings of Interspeech,* pp. 684-688, 2017.

[4] T. F. Quatieri, K. Brady, D. Messing, J. P. Campbell, W. M. Campbell, M. S. Brandstein, *et al.*, "Exploiting nonacoustic sensors for speech encoding," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 14, pp. 533-544, 2006.

[5] R. E. Hillman and D. D. Mehta, "Ambulatory monitoring of daily voice use," *Perspectives on Voice and Voice Disorders,* vol. 21, pp. 56-61, 2011.

[6] I. R. Titze, J. G. Švec, and P. S. Popolo, "Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues," *Journal of Speech, Language, and Hearing Research,* vol. 46, pp. 919-932, 2003.

[7] D. D. Mehta, M. Zañartu, J. H. Van Stan, S. W. Feng, H. A. Cheyne II, and R. E. Hillman, "Smartphone-based detection of voice disorders by long-term monitoring of neck acceleration features," *Proceedings of the IEEE International Conference on Body Sensor Networks,* pp. 1-6, 2013.

[8] 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 Transactions on Biomedical Engineering,* vol. 59, pp. 3090-3096, 2012.

[9] J. H. Van Stan, D. D. Mehta, D. Sternad, R. Petit, and R. E. Hillman, "Ambulatory voice biofeedback: Relative frequency and summary feedback effects on performance and retention of reduced vocal intensity in the daily lives of participants with normal voices," *Journal of Speech, Language, and Hearing Research,* vol. 60, pp. 853-864, 2017.

[10] L. J. Brattain, R. Landman, K. A. Johnson, P. Chwalek, J. Hyman, J. Sharma, *et al.*, "A multimodal sensor system for automated marmoset behavioral analysis," in *IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2016, pp. 254-259.

[11] P. Matychuk, "The role of child-directed speech in language acquisition: a case study," *Language Sciences,* vol. 27, pp. 301-379, 2005.

[12] D. Mehta, J. Van Stan, and R. Hillman, "Relationships between vocal function measures derived from an acoustic microphone and a subglottal neck-surface accelerometer," *IEEE/ACM Transactions on Audio, Speech, and Language Processing,* vol. 24, pp. 659-668, 2016.

[13] 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 Transactions on Audio, Speech, and Language Processing,* vol. 21, pp. 1929-1939, 2013.

[14] V. M. Espinoza, M. Zañartu, J. H. Van Stan, D. D. Mehta, and R. E. Hillman, "Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction," *Journal of Speech, Language, and Hearing Research,* vol. 60, pp. 2159-2169, 2017.