Wireless neck-surface accelerometer and microphone on flex circuit with application to noise-robust monitoring of Lombard speech

Daryush D. Mehta[a,b,c,d], Patrick C. Chwalek[d], Thomas F. Quatieri[d], Laura J. Brattain[d]

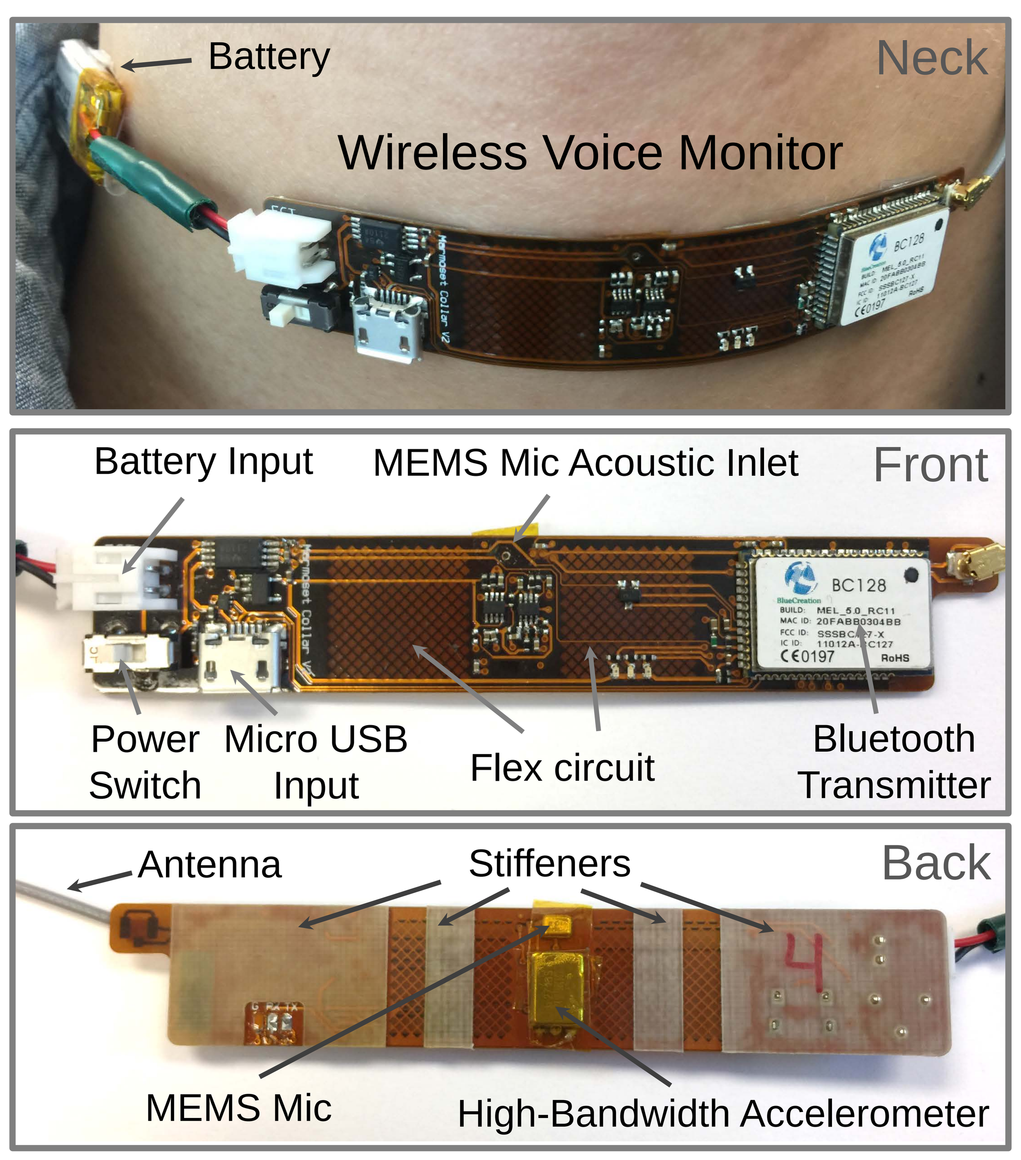
[a]Massachusetts General Hospital, Boston, MA; [b]Harvard Medical School, Boston, MA; [c]MGH Institute of Health Professions, Boston, MA; [d]MIT Lincoln Laboratory, Lexington, MA

Motivation

Ambulatory monitoring of voice and speech characteristics has the potential to provide valuable data for the diagnosis, treatment, and prevention of voice and speech disorders, neurological conditions affecting speech production, and the overall assessment of one's psychological and emotional state. Characterizing **background noise levels and Lombard speech in naturalistic environments** is of critical importance to clinical voice assessment using ambulatory monitoring technology.

Hardware Design

- Novel development of a **lightweight, wireless voice monitor** using Bluetooth technology
- Synchronous recording of **dual-channel data** from an acoustic microphone and a neck-surface accelerometer
- **Flex circuit architecture**
- Full control over hardware and software with access to raw data streams



Two sensors:

- **Single-axis, high-bandwidth accelerometer** (BU-27135, Knowles Electronics, Itasca, IL) placed just above the collarbone and attached to the neck skin
- **Omnidirectional MEMS microphone** (SPA2410LR5H-B, Knowles Electronics) housed adjacent to the accelerometer.

System specifications

Feature	Specification
Sample rate	44.1 kHz (per channel)
Resolution	16 bits
Bandwidth	ACC: 0–5 kHz, MIC: 0–15 kHz
Power consumption	50 mW (transmitting), 18 mW (standby)
Battery life	Up to 8 hours (110 mAh battery)
Weight	4.0 g (12.5 g with 110 mAh battery)
Size	Transmitter: 68 mm x 14.5 mm x 5 mm Receiver: 59 mm x 25 mm x 10 mm
Wireless protocol	Bluetooth 4.0

Study Design

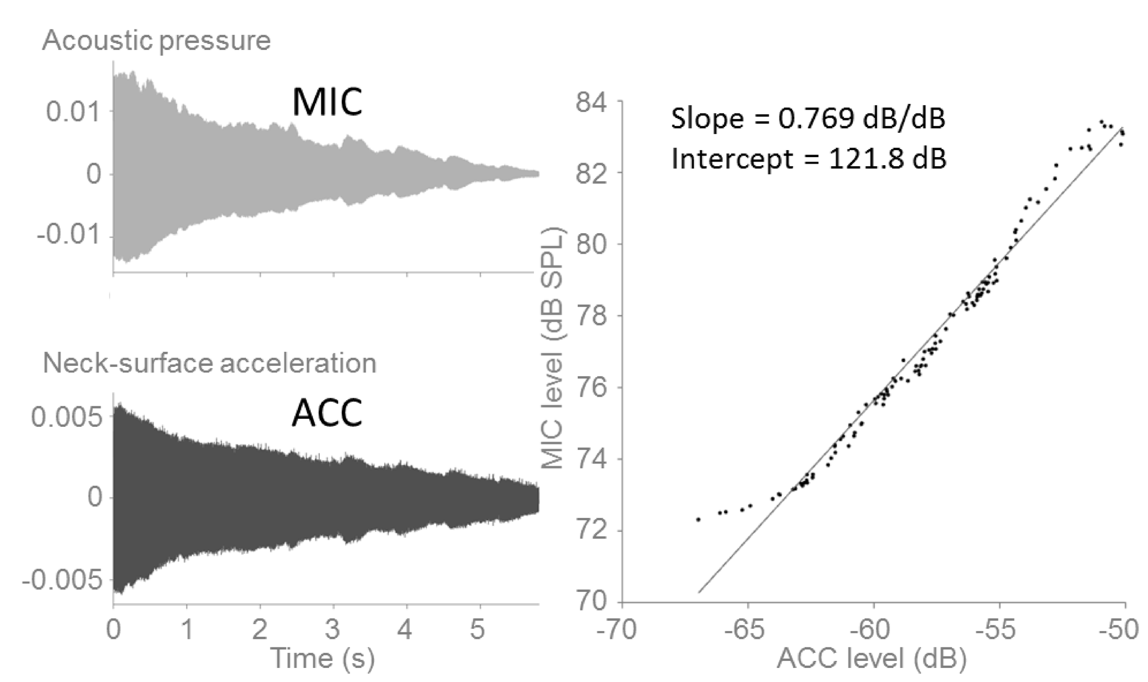
Four adult participants (two male, two female) wore the wireless voice monitor inside an acoustically treated sound booth that contained loudspeakers that allowed for the **simulation of ambient acoustic stimuli** at varying calibrated sound pressure levels. Each participant performed the following speech tasks:

1) /a/ vowel starting at a loud intensity and gradually decreasing to a soft level
2) Phonetically balanced Rainbow Passage

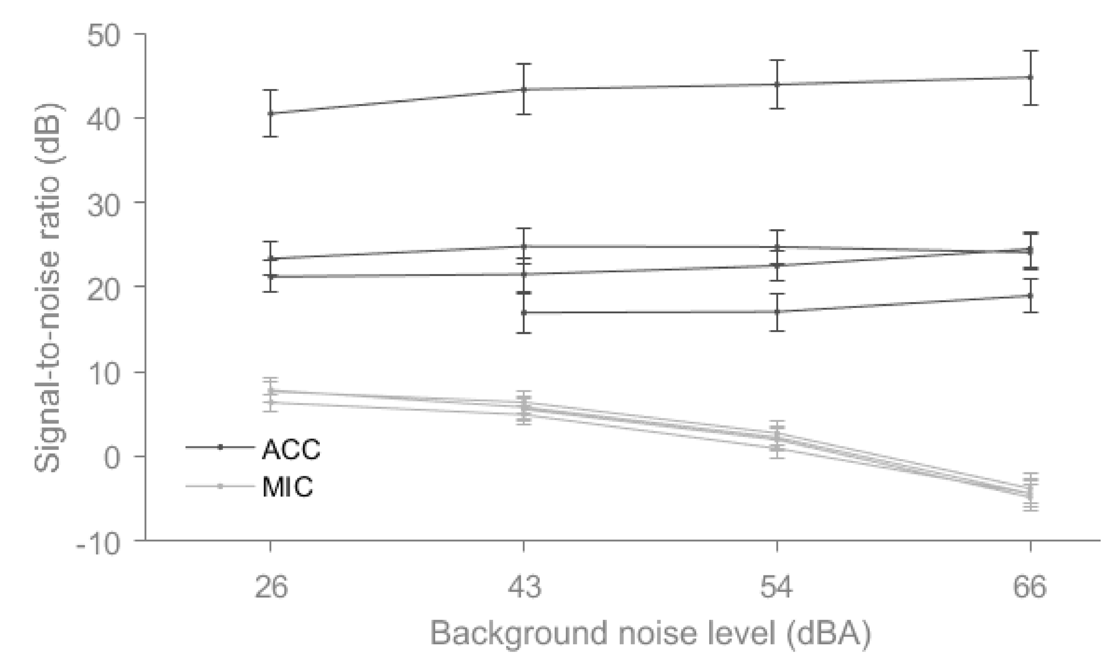
Four different levels of the same background noise stimulus (helicopter rotors):

1) Quiet – 26 dBA
2) Mild – 43 dBA
3) Moderate – 54 dBA
4) Loud – 66 dBA

Results: Robustness of Accelerometer to Ambient Noise

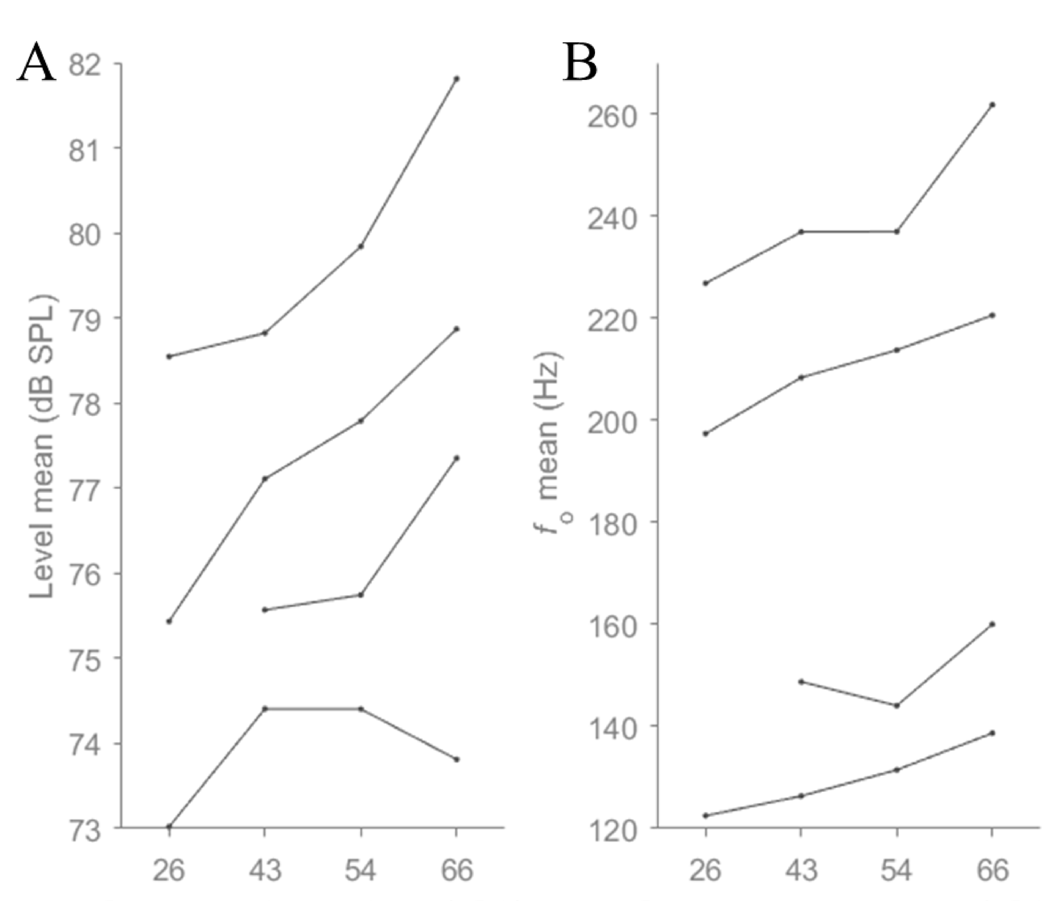


Participant-specific calibration of accelerometer (ACC) signal level to microphone (MIC) sound pressure level using the loud-to-soft vowel task.



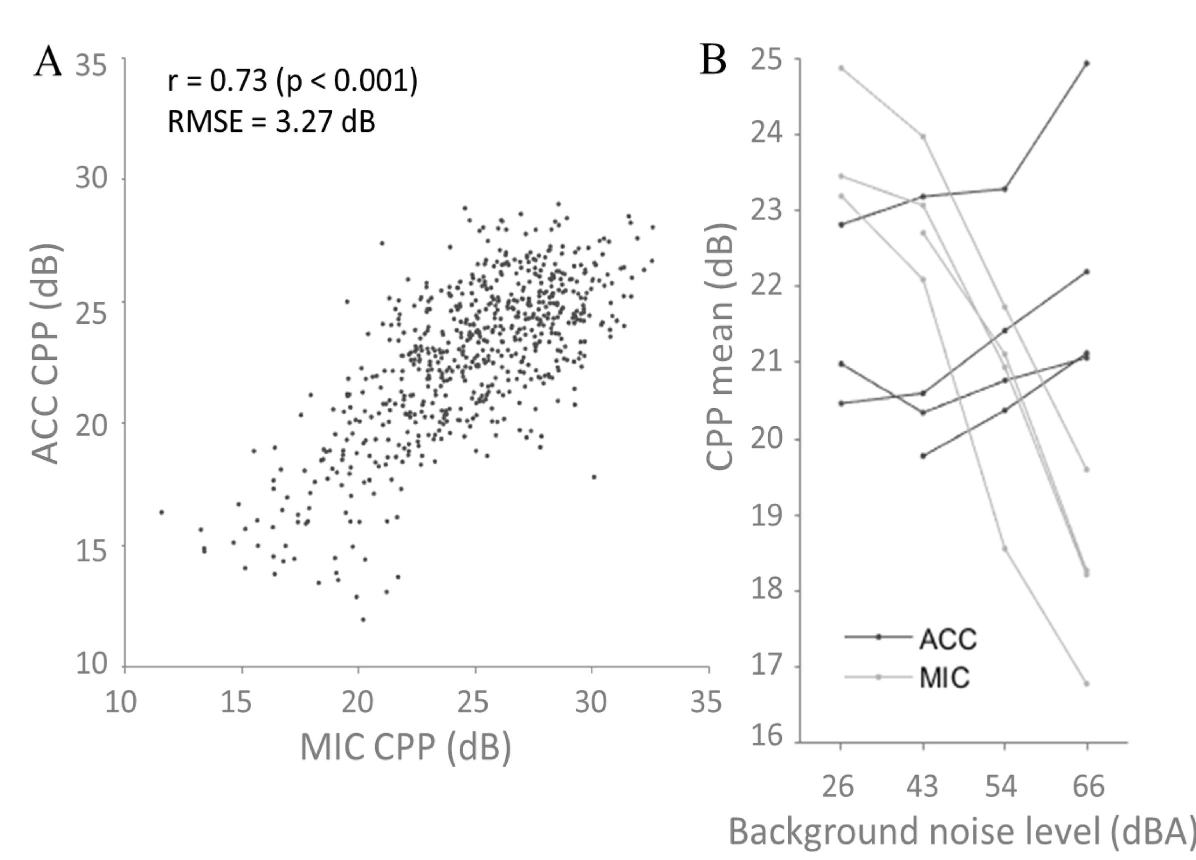
Comparison of signal-to-noise ratio (SNR) of the ACC and MIC signals for the Rainbow Passage. Error bars: ± 1 std. dev.

Results: Lombard Speech Effects during Connected Speech



Level and fundamental frequency increases

(A) Mean sound pressure level increases
(B) Mean fundamental frequency (f_o) increases



Mean cepstral peak prominence (CPP) increases

(A) Correlation between frame-level MIC- and ACC-based CPP (quiet) Pearson's r and root-mean-square error (RMSE) shown.
(B) Lombard effect only observed in the ACC signal

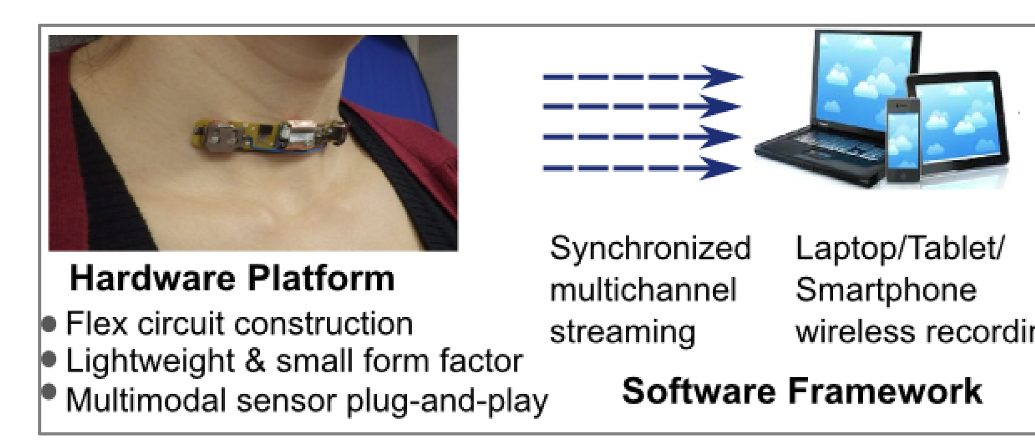
Discussion

Future work

- **Ambulatory tracking of everyday verbal communication** as individuals go about their typical daily activities.
- Efforts to develop a custom wireless solution have been motivated by experience demonstrating that **patient compliance improves when technology is easy to use** and less cumbersome.
- Future device development can also take advantage of the modularity of the system to add **additional sensors and real-time processing of voice features that can provide user biofeedback** via mobile devices such as smartphones and smartwatches.

Summary of results

- **Accelerometer (ACC)-based SNR remained stable across all background noise levels** when compared with the decreasing values for MIC-based SNR.
- Estimates of voice SPL may be better obtained using the ACC signal as compared to the MIC signal in naturalistic environments that exhibit varying levels of background acoustic noise.
- **Participant-specific SPL mapping required in a quiet setting** to be applied to ACC signal levels in noisy settings.
- **ACC-based estimates of CPP** can act as noise-robust measures of overall voice quality.



Conclusion

A **new wireless voice monitor has been developed that uses Bluetooth technology and a wearable, flexible circuit**. Synchronized data streaming from both acoustic and neck-surface sensors makes it feasible to compute complementary acoustic and non-acoustic speech/voice features. The microphone also provides critical information related to environment noise levels that is important to collect in real-world conditions. **Characterizing Lombard speech**—a speaker's vocal reaction to varying ambient noise levels—is enabled with synchronous ACC and MIC recordings.

Acknowledgments

Special thanks to Prof. Robert Desimone, Prof. Guoping Feng, and Dr. Charles Jennings at the McGovern Institute for Brain Research at MIT, Dr. Rogier Landman at the MIT/Harvard Broad Institute, and Mr. Kerry Johnson, Mr. Tejash Patel, and Dr. Christopher Smalt at MIT Lincoln Laboratory for their generous support and help.

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. This work was also supported by the MIT McGovern Institute Neurotechnology Program.