Article

Wearable Impedance-Matched Noise Canceling Sensor for Voice Pickup [†]

Hee Yun Suh, Helena Hahn and James West *

Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218, USA; email1@email.com (H.Y.S.); email2@email.com (H.H.)

* Correspondence: jimwest@jhu.edu
† Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–30 November 2023; Available online: https://ecsa-10.sciforum.net/.

Abstract: Communicating under extreme noise conditions remains challenging in spite of higher-order noise-canceling microphones, throat microphones, and signal processing. Both natural and man-made background ambient noise can disturb the conveyance of information because of high noise levels. Noise cancellation, which is used frequently in audio technology, has limits in noise reduction and does not guarantee clear vocal pickup in these severe situations. A contact microphone that is attached directly to the medium of interest has the potential to pick up vocal signals with reduced noise. In this study, an electrostatic transducer with an elastomer layer that is impedance-matched to the human body is used to pick up speech sounds through constant contact on the chin and cheek. By attaching the wearable device directly to the skin, the medium of air is bypassed, and airborne noise is passively canceled. Because of the acoustic impedance-matched layer, the sensor is more sensitive to low frequencies under 500 Hz, so frequency equalization was implemented to flatten the frequency response throughout the vocal range. The perceptual evaluation of speech quality (PESQ) scores of the wearable device with equalization averaged around 2.6 on a scale from –0.5 to 4.5. Speech recordings were also collected in a noise field of 85 dB, and the performance was compared to a cardioid lapel mic, a cardioid dynamic mic, and an omnidirectional condenser mic. The recordings revealed a significantly reduced presence of white noise in the contact sensor. This study provides preliminary results that show potential vocal applications for a wearable impedance-matched sensor.

Keywords: wearable device; acoustic noise; transducers; speech communication; microphones; noise cancellation

1. Introduction

Extreme noise conditions such as construction and heavy traffic can reduce the quality of communication. Active noise cancellation, piezoelectric throat microphones, and signal processing techniques are methods used to improve the conveyance of information in the presence of ambient noise. Efforts in active noise cancellation and adaptive filtering algorithms provide attenuation of about 20–30 dB [1,2]. This may not be sufficient for optimal communication and voice pickup in environments with high noise levels. Various higher-order microphones have been developed for improved directionality, but many higher-order microphones also have high noise sensitivity [3].

A recently developed electrostatic transducer has the potential to reduce high noise levels while picking up sounds. The sensor's elastomer layer is impedance-matched to the skin, the medium of interest [4,5]. By attaching the device to the skin, the medium of air is bypassed, so the transducer passively rejects airborne noise while reducing the loss of signal energy [5]. The impedance-matched sensor has been implemented in a wide range of settings such as musical acoustics [6] and body sound monitoring [5]. When the transducer is placed on areas with high vocal vibration, such as under the chin or on the cheek, it can be used as a wearable sensor with high noise-cancellation abilities for voice pickup. This

paper focuses on enhancing the speech-pickup abilities of the acoustic impedance-matched transducer and comparing it to more widely used microphones to demonstrate the sensor's potential in vocal applications.

2. Materials and Methods

2.1. Impedance-Matched Transducer

The electrostatic transducer was created with a tuned elastomer layer with coated microstructures and a charged fluorinated ethylene propylene (FEP) film as seen in Figure 1. Corona charging was used to charge the FEP film, and the layers were encased with shielding to create a thin shape (Figure 2a,b).
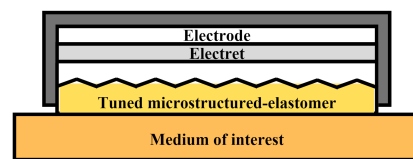

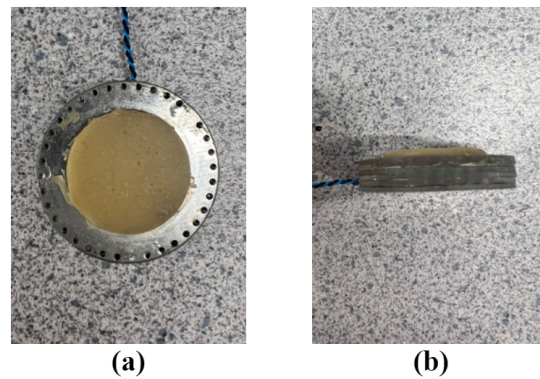**Figure 1.** Inner layers of transducer.

(a)   **(b)**

Figure 2. (**a**) Bottom view of the transducer. (**b**) Side view of the transducer.

2.2. Experimental Setup

Medical tape (3M Tegaderm™) was used to adhere the device to the cheek and under the chin at positions shown in Figure 3. For comparison, three conventionally used microphones were selected: a cardioid lapel microphone (Audio-Technica AT898 Lavalier Mic), a cardioid dynamic microphone (Sennheiser e835 Dynamic Mic), and an omnidirectional condenser microphone (Blue Yeti Pro Mic on omnidirectional mode). A Focusrite Scarlett 2i4 Audio Interface collected the output of the transducer and microphones, and Audacity was used to record the audio. Phantom power of 48 V was supplied by the audio interface for all recordings, and the gain was adjusted to avoid clipping.
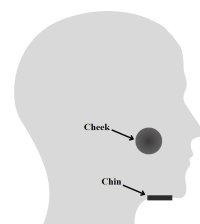

Figure 3. Position of the transducer when taped to the cheek and under the chin on the subject.

The microphones were positioned based on polar patterns. The cardioid dynamic mic was placed on a tabletop microphone holder at a 45-degree angle downward and about 1 inch away from the subject's mouth to ensure maximum pickup. The omnidirectional

condenser mic is best placed 4–10 inches away from the source, so it was propped upward and placed 8 inches away from the subject's mouth. The cardioid lapel mic was attached facing upward to the upper chest area of the subject with a magnetic clip.

2.3. Speech Quality

To measure the quality of the speech recorded by the transducer, the first list in the Harvard sentences was used [7]. Each list in the Harvard sentences is phonetically balanced and widely used for speech quality measurements. The subject recorded all 10 sentences from List 1, and additional samples included a tongue-twister and counting from 1–10 (Table 1). The transducer, lapel, and omnidirectional mic were each simultaneously recorded with the dynamic mic, which was set as the reference because of its robustness.

Table 1. Sentences recorded for speech quality experiment.

Type	Sentence
Counting	One, two, three, four, five, six, seven, eight, nine, ten.
Tongue-twister	She sells seashells by the seashore.
Harvard sentences List 1	1. The birch canoe slid on the smooth planks. 2. Glue the sheet to the dark blue background. 3. It's easy to tell the depth of a well. 4. These days a chicken leg is a rare dish. 5. Rice is often served in round bowls. 6. The juice of lemons makes fine punch. 7. The box was thrown beside the parked truck. 8. The hogs were fed chopped corn and garbage. 9. Four hours of steady work faced us. 10. A large size in stockings is hard to sell.

2.3.1. Post-Processing

Key phonetic features exist in the frequency range up to 6–8 kHz, and higher frequencies can also add spectral information [8]. The acoustic transducer is sensitive to lower frequencies below 500 Hz, so post-processing was needed to enhance the speech recordings. Lowpass filtering at 40 Hz was done to remove unwanted noise from the impedance-matched transducer recordings. The recordings from the transducer were then amplified by 8 dB to match the volume of the other microphones.

To flatten the frequency response of the transducer's output, commercial equalization software (Logic Pro) was used to match the frequency response of the transducer recording to that of the dynamic mic recording. Equalization had a cutoff frequency at around 1300 Hz for the cheek-positioned recordings and a cutoff frequency at around 1900 Hz for the chin-positioned recordings to minimize high-frequency noise. These cutoffs were determined by maximizing the perceptual evaluation of speech quality (PESQ) [9].

2.3.2. Speech Quality Measurement

To quantify the quality of the speech, the PESQ score was calculated. PESQ takes into consideration noise and audio distortion [9]. The scores range from -0.5 to 4.5 with scores between 2 and 3 needing moderate effort to understand, and scores 3 and above needing less effort to understand.

2.4. Noise Cancellation

A noise field was created using two speakers (Yamaha HS8) placed at two opposing corners of a sound booth. White noise generated in Audacity was played through both speakers synchronously, and the decibel level was measured using a sound level meter (Martel 322). The decibel levels were measured in dBA with the meter placed at the site of the recording equipment. The subject sat between the two speakers with the transducers, and the microphones were positioned similarly to the description above (Figure 4). An '/a/'

sound was recorded at noise levels of 60 dB to 85 dB in 5 dB increments, and a reference recording was done without noise.
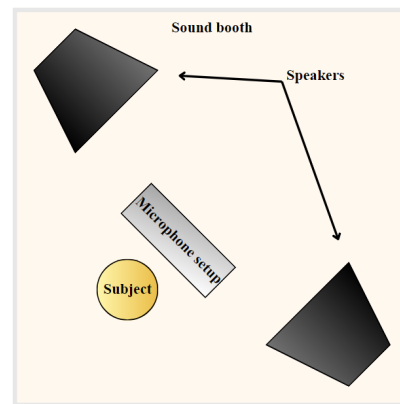

Figure 4. Sound booth set up for noise cancellation experiment.

3. Results

The speech quality and noise cancellation results are summarized below.

3.1. Speech Quality

Table 2 presents the PESQ scores of the impedance-matched transducer and comparison microphones. The PESQ score increased for all sentences after post-processing for the acoustic transducer. The average post-processed transducer score is about 2.59 and about 0.5 less than the average score of the other two microphones. Despite containing more artifacts, such as buzzing, the cheek position had a slightly higher post-processed PESQ score than of the cheek position.

Table 2. Table of PESQ scores for each transducer position and microphone. The dynamic mic was used as the reference.

Sentence	Transducer (Cheek) Original	Transducer (Chin) Original	Transducer (Cheek) Post-Processed	Transducer (Chin) Post-Processed	Lapel Mic	Omnidirectional Mic
Counting	1.4783	2.1966	3.0238	3.0599	3.4527	3.4018
Tongue-Twister	1.8844	2.0097	2.3804	2.2728	3.1010	2.9669
1 *	1.6549	1.4344	2.6274	2.4440	3.1259	3.0640
2	1.4278	1.3410	2.2930	2.9153	3.2753	3.2726
3	1.8041	1.6775	2.8402	2.6124	3.0955	3.0113
4	1.6774	1.7726	2.3861	2.6759	3.1076	3.0472
5	1.6883	1.6720	2.3723	2.3198	3.0204	2.9092
6	1.8740	1.5036	2.8963	2.3081	3.0204	2.9929
7	1.7565	1.8316	2.6523	2.4899	3.2521	3.0875
8	1.8423	2.1095	2.9682	2.6612	3.2204	3.0691
9	1.7113	1.6813	2.5233	2.4539	3.2069	3.2400
10	1.8941	1.8283	2.5802	2.4474	3.1858	2.8575
Average	1.7245	1.7548	2.6286	2.5551	3.1720	3.0676

* All numbered sentences are from List 1 of the Harvard Sentences.

3.2. Noise Cancellation

Spectrograms of the signals (Figure 5) show reduced noise in the transducer recordings compared to the other microphones. The white noise visibly appears throughout the frequency range in the audio signals for the comparison microphones at 85 dB. The transducer audio signals have very little visible white noise, and they are similar throughout

different noise-level environments. The white noise is also audibly significant in the three microphone recordings compared to the transducer recordings.
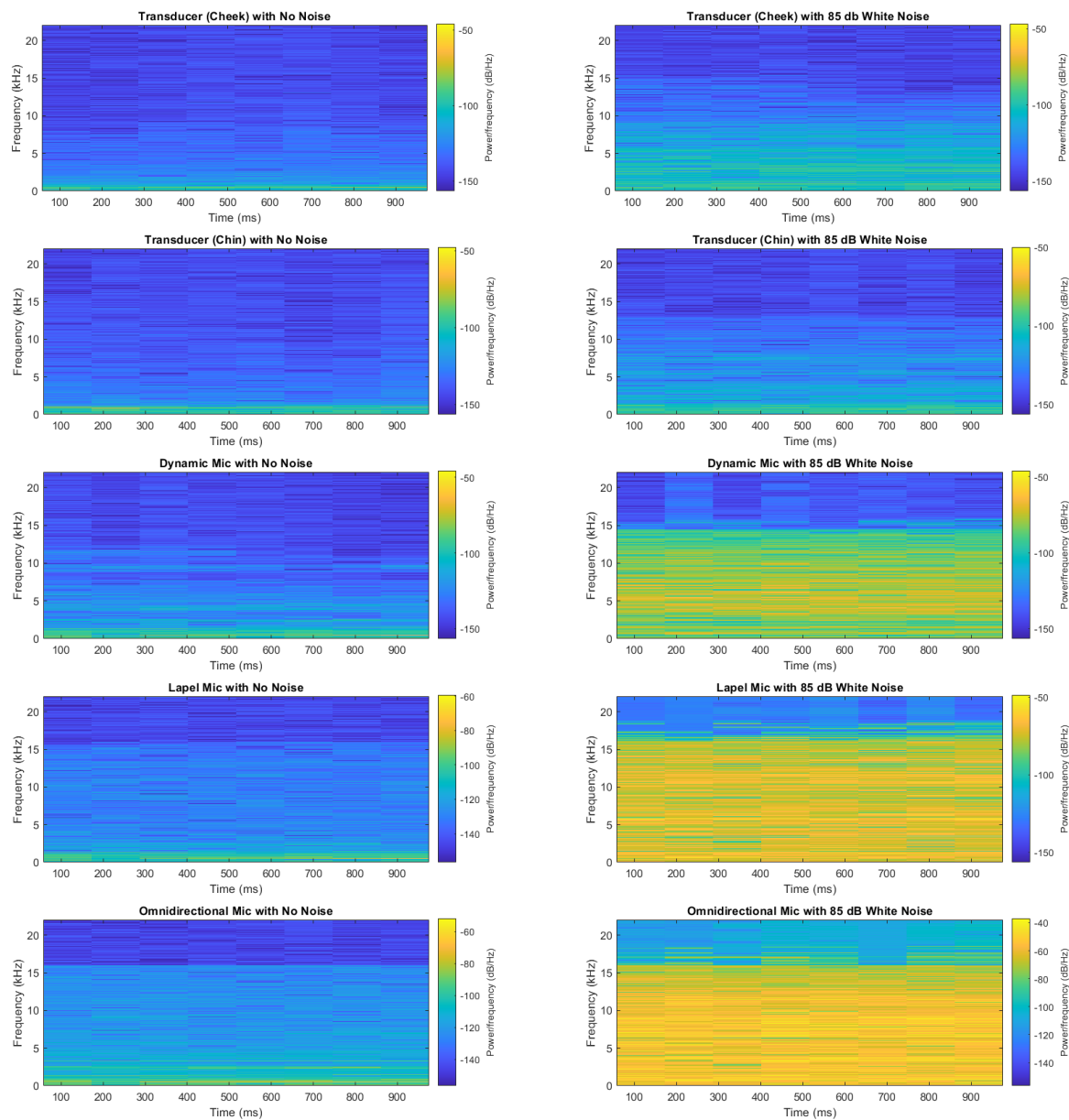

Figure 5. Spectrograms of the recordings at no noise and 85 dB noise using Hamming window and 50% overlap.

4. Conclusions

The current paper provides preliminary results on the vocal pickup abilities of an acoustic impedance-matched transducer. The transducer captured similar audio quality at the cheek and chin positions. Equalization improved the speech quality of the transducer recordings, increasing the PESQ score to an intelligible level of 2.6 from the original 1.7. The speech quality may reach a similar standing to other microphones if post-processing methods and additional transducer tuning techniques are further investigated. For noise cancellation, the transducer proved to have superior noise reduction capabilities in comparison to three different microphones. Little noise was detected even at loud noise levels of 85 dB. Future comparison with contact microphones such as throat microphones may prove to be helpful. The impedance-matched sensor demonstrates potential as a wearable

noise-canceling contact microphone in vocal applications particularly for extreme noise conditions.

References

1. Boll, S.; Pulsipher, D. Suppression of acoustic noise in speech using two microphone adaptive noise cancellation. *IEEE Trans. Acoust. Speech Signal Process.* **1980**, *28*, 752–753. https://doi.org/10.1109/TASSP.1980.1163472.
2. Dixit, S.; Nagaria, D. LMS Adaptive Filters for Noise Cancellation: A Review. *Int. J. Electr. Comput. Eng. (IJECE)* **2017**, *7*, 2520–2529. https://doi.org/10.11591/ijece.v7i5.pp2520-2529.
3. De Sena, E.; Hacihabiboglu, H.; Cvetkovic, Z. On the Design and Implementation of Higher Order Differential Microphones. *IEEE Trans. Audio Speech Language Process.* **2012**, *20*, 162–174. https://doi.org/10.1109/TASL.2011.2159204.
4. Rennoll, V.; McLane, I.M.; Eisape, A.; Elhilali, M.; West, J. Evaluating the impact of acoustic impedance matching on the airborne noise rejection and sensitivity of an electrostatic transducer. *J. Acoust. Soc. Am.* **2021**, *149*. https://doi.org/10.1121/10.0004413.
5. Rennoll, V.; McLane, I.M.; Eisape, A.; Grant, D.; Hahn, H.; Elhilali, M.; West, J. Electrostatic Acoustic Sensor with an Impedance-Matched Diaphragm Characterized for Body Sound Monitoring. *ACS Appl. Bio Mater.* **2023**, *6*, 3241–3256. https://doi.org/10.1021/acsabm.3c00359.
6. Rennoll, V.; McLane, I.M.; Eisape, A.; Grant, D.; Betz, C.; Chen, X.; Gebhart, M.; Hahn, H.; Kartub, S.; Lehr, B.; et al. Project-based learning through sensor characterization in a musical acoustics course, *J. Acoust. Soc. Am.* **2022**, *152*, 1932–1941. https://doi.org/10.1121/10.0014171.
7. IEEE Recommended Practice for Speech Quality Measurements. *IEEE Trans. Audio Electroacoust.* **1969**, *17*, 225–246. https://doi.org/ 10.1109/TAU.1969.1162058.
8. Trine, A.; Monson, B.B. Extended High Frequencies Provide Both Spectral and Temporal Information to Improve Speech-in-Speech Recognition. *Trends Hear.* **2024**, *24*. https://doi.org/10.1177/2331216520980299.
9. Rix, A.W.; Beerends, J.G.; Hollier, M.P.; Hekstra, A.P. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal, Salt Lake City, UT, USA, 7–11 May 2001; Volume 2, pp. 749-752. https://doi.org/10.1109/ICASSP.2001.941023.