# User Authentication Method for Hearables Using Sound Leakage Signals


Takashi Amesaka
Keio University
amesaka@keio.jp

Hiroki Watanabe
Hokkaido University
hiroki.watanabe@ist.hokudai.ac.jp

Masanori Sugimoto
Hokkaido University
sugi@ist.hokudai.ac.jp

Yuta Sugiura
Keio University
sugiura@keio.jp

Buntarou Shizuki
University of Tsukuba
shizuki@cs.tsukuba.ac.jp



## ABSTRACT

We propose a novel biometric authentication method that leverages sound leakage signals from hearables that are captured by an external microphone. A sweep signal is played from hearables, and sound leakage is recorded using an external microphone. This sound leakage signal represents the acoustic characteristics of the ear canal, auricle, or hand. Then, our system analyzes the echoes and authenticates the user. The proposed method is highly adaptable to hearables because it leverages widely available sensors, such as speakers and external microphones. In addition, the proposed method has the potential to be used in combination with existing methods. In this study, we investigate the characteristics of sound leakage signals using an experimental model and measure the authentication performance of our method using acoustic data from 16 people. The results show that the balanced accuracy (BAC) scores were in the range of 87.0%–96.7% in several scenarios.



## CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**.

## KEYWORDS

hearables, authentication, active acoustic sensing, sound leakage, earables

**ACM Reference Format:**
Takashi Amesaka, Hiroki Watanabe, Masanori Sugimoto, Yuta Sugiura, and Buntarou Shizuki. 2023. User Authentication Method for Hearables Using Sound Leakage Signals. In *Proceedings of the 2023 International Symposium on Wearable Computers (ISWC '23), October 08–12, 2023, Cancun, Quintana Roo, Mexico.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3594738.3611376


## 1 INTRODUCTION

Earphone-type wearable devices, also known as "hearables," are gaining attention as a new platform with various functions [17], such as music, communication, vitals sensing for health management [8, 11], and voice assistance for daily life activities. It is anticipated that some applications of hearables will require the authentication of the wearer because they may access personal information, such as messages, calendars, and biometric information [6]. Moreover, with the potential development of hearables authentication technology, offering a service that enables users to procure goods or digital content from e-commerce platforms solely through a device's voice assistant may be feasible. Therefore, the development of such an authentication method is deemed crucial for advancing the future of hearables.

However, the application of existing authentication methods used in smartwatches or smartphones to hearables poses challenges. Hearables lack touch panels and have limitations in terms of using fingerprint sensors. Voiceprint authentication is a potential solution; however, it poses challenges in noisy environments and public places. To address these challenges, previous studies have proposed authentication methods for hearables [1, 3–5, 7, 9, 10, 15, 19, 20, 22]. These methods utilize inward-facing microphones to acquire acoustic features of the ear canal, teeth, gait, and heartbeat for authentication.

In this study, we propose a novel biometric authentication method that leverages sound leakage signals from hearables captured by an external microphone (Figure 1). As the shapes of the ear canal and auricle vary from person to person, the acoustic characteristics of these sound leakage signals also differ between individuals. We exploit these differences in acoustic characteristics for wearer authentication. This is the first method that focuses on using the external microphone, which has not been tested in previous research.

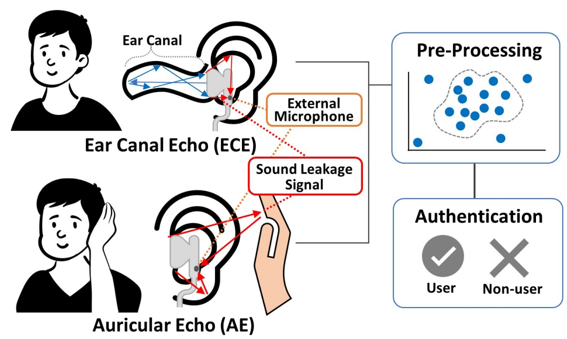


**Figure 1: Flow of the authentication system used in the proposed method.**



*ISWC '23, October 08–12, 2023, Cancun, Quintana Roo, Mexico*



ACM ISBN 979-8-4007-0199-3/23/10…$15.00
https://doi.org/10.1145/3594738.3611376

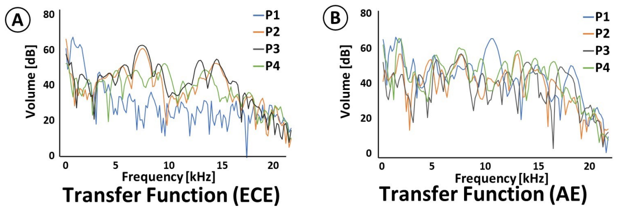


**Figure 2: Transfer function for each participant. A: ear canal echo, B: auricular echo**

The proposed method is highly adaptable to hearables because it leverages widely available sensors, such as speakers and external microphones. In addition, the proposed method can be used both independently and cooperatively with authentication methods that use an inward-facing microphone. This has the potential to enhance the robustness of hearables authentication.

We investigated sound leakage signals and their acoustic characteristics; then, we conducted authentication experiments with 16 participants in noisy and walking environments. The results show that the balanced accuracy (BAC) scores, which represent the mean of the false acceptance rate (FAR) and the false rejection rate (FRR), were in the range of 87.0%–96.7%.

## 2 RELATEDWORK

In this section, we describe previous research on personal authentication methods for hearables and research that uses the individuality of the auricle's shape.

### 2.1 Authentication Methods for Hearables

EarEcho [10] and EarDynamic [20] are personal authentication methods that leverage the acoustic characteristics of the ear canal measured using an inward-facing microphone. ToothSonic [19] and TeethPass [22] use an inward-facing microphone to record the sound produced when the user's teeth mesh for personal authentication. HeartPrint [3] uses an inward-facing microphone built into hearables to record the wearer's heartbeat for personal authentication. EarGate [7] uses the inward-facing microphone of hearables to record the sound generated by the wearer's body while walking and acquires the user's gait data; then, it performs personal authentication based on the gait data obtained over several steps. Vocal Resonance [15] and Voice In Ear [9] also perform personal authentication using the voice sound acquired from the inward-facing microphone of hearables.

The abovementioned methods use inward-facing microphones to acquire the echoes of the ear canal, sounds propagating inside the body, and the voice of the user. However, inward-facing microphones are only installed in some hearables, which are canal-type devices, on the market because they are used for high-performance noise cancellations. In contrast, focusing on signals leaking outward, our proposed method uses external microphones to acquire echoes from the ear canal and the auricular portion of the ear for personal authentication.

### 2.2 Individuality of the Auricle

The auricle efficiently collects ambient sound and transmits it to the ear canal [13, 16], which differs from person to person; therefore, some studies use the individuality of the auricle's shape for personal identification/authentication. Holz et al. [12] proposed a personal identification system that uses the touch screen of a smartphone as a sensor to acquire spatial feature points of the auricle when the smartphone is placed on the ear. Itani et al. [14] proposed a system for personal authentication through the acquisition of the auricular transfer function by placing a smartphone-type device against the ear. Sulavko et al. [18] acquired the acoustic characteristics of the auricle and ear canal using headphones and investigated the individuality of the auricle and ear canal.

Similar to these methods, our proposed method aims at personal identification by using the individuality of the auricle's shape. Specifically, our proposed method focuses on sound leakage signals and hand-covering movements to acquire the acoustic characteristics of the auricle. In addition, it uses the speaker and external microphone, which are commonly installed in earphone-type devices; the devices used in previous studies were not earphone-type devices.

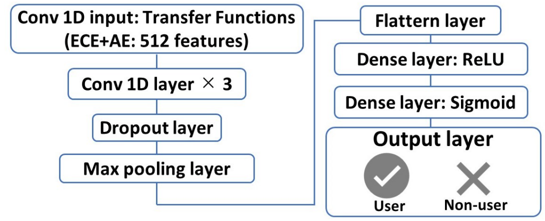


**Figure 3: The architecture of the convolutional neural network model of the proposed method.**

## 3 PROPOSED METHOD

The overall flow of our authentication method is shown in Figure 1. The system records echoes from two different user states; then, using the obtained acoustic data, it performs pre-processing and generates a machine-learning model to authenticate the user.

### 3.1 Recording of Sound Leakage Signals

In the recording phase, the first measurement is made when the user wears the device. The device transmits an up-sweep signal as a measurement signal. The measurement signal reverberates in the ear canal and the auricular geometry around the device, and some of the signals leak to the outside. This sound leakage is recorded by an external microphone as an ear canal echo. Then, the second measurement is made when the user covers the entire ear with his/her hand; the sound leakage signal reverberates in the space formed by the hand and the auricle. This echo is recorded by the external microphone as an auricular echo.

### 3.2 Pre-processing

In the pre-processing phase, a transfer function with a high signal-noise-ratio (SNR) is acquired from the echoes. The synchronous addition and cut-out processes are performed to obtain the high-SNR transfer function of 256 dimensions. The transfer function of both echoes for four participants are shown in Figure 2, which shows that the transfer functions obtained are different for each participant.

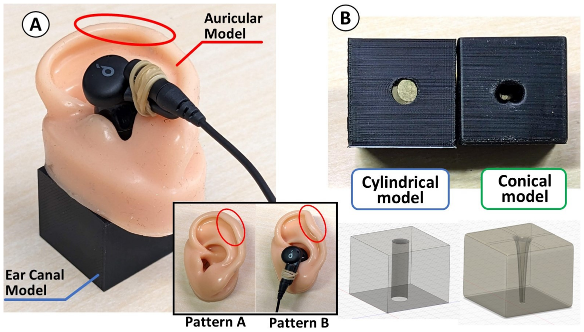


**Figure 4: Physical model for testing the acoustic characteristics of the sound leakage signal. A: the entire model, B: the ear canal models.**

In addition, outlier removal is performed on the obtained data to avoid providing the authentication model with data that are likely to be misidentified. First, the outlier detection model is created using the isolation forest algorithm. We use the created model to detect and exclude outliers from the training test data.

## 3.3 Authentication Model

In the authentication model phase, a two-class convolutional neural network (CNN) model is used for user authentication. As shown in Figure 3, we use a one-dimensional CNN model with three hidden convolutional layers followed by a dropout layer, a max pooling layer, a flatten layer, two dense layers, and an output layer. The convolution layer has 64, 128, and 256 filters. The activation functions for each dense layer were rectified linear unit (ReLU) and sigmoid functions.

# 4 ACOUSTIC CHARACTERISTICS OF SOUND LEAKAGE SIGNALS

To the best of our knowledge, there are no authentication systems that utilize ear canals and auricular echoes obtained from sound leakage signals. Therefore, we performed a quantitative study using a physical ear model to understand the properties of the sound leakage signal. We used a physical model consisting of ear canal and auricular models, as shown in Figure 4A. For the ear canal model, two blocks with cylindrical and conical cavities were created using a 3D printer (Figure 4B); the eardrum was reproduced by sealing one side of the cavity with paper tape. For the auricular models, an off-the-shelf silicon ear model (TOYMYTOY) was used; it was used as is (Pattern A) and in a deformed state (Pattern B) by passing a wire through the red circled area in Figure 4A. By combining each model, the relationship between the shape change in each part and the sound leakage signal was investigated.

The frequency responses of the ear canal and auricle are summarised in Figures 5. In this investigation, we obtained the frequency responses of the cylindrical and conical ear canal models with the auricle in Pattern A (Ear Canal Echo Test). In addition, we obtained the frequency responses of the cylindrical ear canal model with the auricles in Pattern A and Pattern B (Auricular Echo Test). Furthermore, we obtained the frequency responses of the cylindrical ear canal model with the auricles in Pattern A and Pattern B when the entire ear was covered with a paper cup (Auricular Echo Test with Paper Cup).

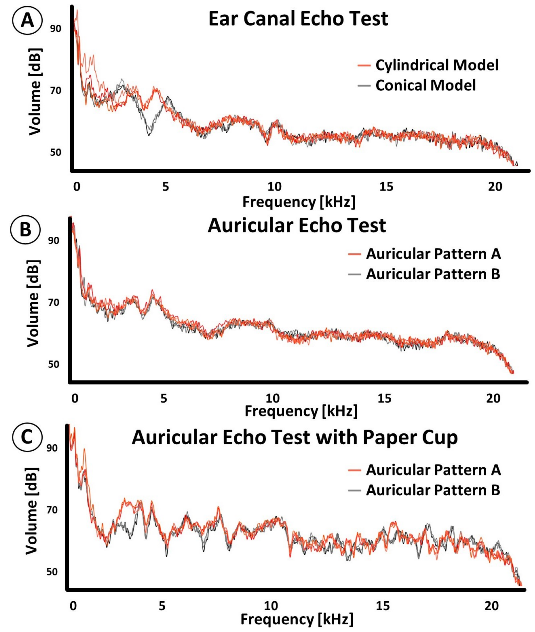


**Figure 5: Frequency response in each test. A: ear canal echo test, B: auricular echo test, C: auricular echo test (paper cup).**

Figure 5A shows that the frequency response obtained changes by up to 10dB (4 kHz to 5 kHz) for different ear canal models. As shown in Figure 5B, the auricular deformation in the red-circled area did not have a significant effect when the sound leakage signal did not reverberate. However, when the entire ear was covered with a paper cup, the shape of the auricle was found to affect the reverberation (Figure 5C). These investigations suggest that the acoustic properties of the ear canal's shape predominate in the sound leakage signal when the hand is not covered, whereas the acoustic properties of the auricular shape predominate when the entire ear is covered by the hand.

# 5 EVALUATION

We evaluated the baseline authentication performance of the proposed method by creating a classification model using training data. In addition, we also investigated the authentication performance using the test data obtained in noisy environments and during walking.

## 5.1 Experimental Setup

Figure 6 shows the device used in the experiment. Since microphones in commercially available earphone-type devices are internally low-pass filtered, which limits the acquisition of high-frequency signals, we attached another microphone to a commercial earphone to investigate the actual performance of the proposed method. The earphones were Anker Soundcore Life Note 3S, and the microphone was Audio Technica AT9904. The acoustic signals acquired from the microphone were converted from analog to digital using ZOOM's F1 audio interface and stored in a PC (ASUS: ROG Flow X13). The signal was an up-sweep signal with a frequency

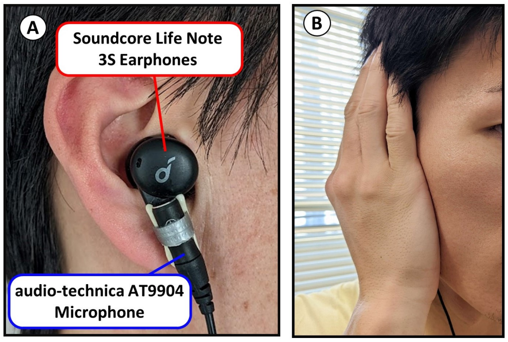


**Figure 6: Experimental device (A) and auricular echo acquisition (B).**

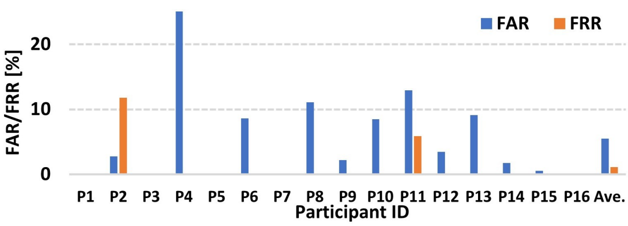


**Figure 7: The false acceptance rate (FAR) and the false rejection rate (FRR) for each participant.**

transition from 0 Hz to 22 kHz. The signal length was 2,048 samples. The sampling rate for the playback/recording of the acoustic signal was 44.1 kHz, and the quantization was 16-bit. The volume was set at approximately 80 dB at the tympanic membrane. This is the same noise level as that in an underground car or aircraft and is considered safe for hearing health when used for short periods [21].

## 5.2 Data Collection

We collected ear canal and auricular echoes from 16 participants (20–29 years old, six females and ten males). All participants were right-handed. The experiment was conducted in a laboratory. The participants were asked to place the measurement device on their right ear while sitting on a chair. We asked the participants to remove their earrings, rings, and glasses, and if their hair covered their ears, we asked them to use clips or hair ties to keep their ears uncovered. Then, we transmitted the measurement signal continuously and recorded it for five seconds to acquire an ear canal echo. Next, we asked the participants to cover their entire right ear with their right hand, as shown in Figure 6B. During this time, we transmitted the measurement signal continuously and recorded it for five seconds to acquire an auricular echo. We defined these measurements as one round and asked the participants to take the device off and put it back in place between each round. In total, we acquired 15 rounds of data from each participant and used them as training data.

## 5.3 Baseline Authentication Performance

To investigate the baseline authentication performance of the proposed method, we evaluated the accuracy of leave-one-round-out cross-validation learning using training data. Specifically, one user and three non-users were selected from the 16 participants; 14

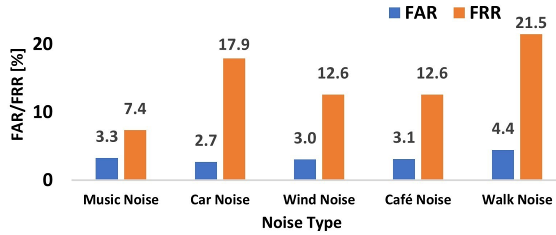


**Figure 8: Average FAR and average FRR for all participants for each noise.**

rounds of user data and data from the remaining 12 participants were used as training data to create a classification model. Since these training data were unbalanced, the number of data for each label was made the same by downsampling. For the test data, we used the remaining one round of user data and one round of non-user data.

The FAR and FRR for each participant are summarized in Figure 7. The average FAR and FRR for all participants were 5.5% and 1.1%, respectively. The BAC score, which represents the mean of the FAR and FRR, averaged 96.7%.

## 5.4 Impact of External Noise

We investigated the impact of external noise and walking scenario on our system. In the noise impact study, we evaluated the authentication performance of test data containing four types of noise sounds (music, car, window, and cafeteria). Noise-containing data were acquired while each noise sound was played from a loudspeaker in front of the participants. The noise level near the participants was approximately 50–60 dB. In the walking impact study, we evaluated the authentication performance of the test data while the participants were walking.

For each type of noise and walking, the average FAR and FRR for all participants are summarized in Figure 8. The BAC scores for each type of noise were 94.7%, 89.7%, 92.2%, 91.6%, and 87.0%, respectively.

# 6 DISCUSSION

EarEcho [10] shows a BAC score of 94.5%–97.5% using the acoustic properties of the ear canal, and the authentication performance of our system is not comparable to it. In the future, detailed acoustic characteristics and effective frequency bands of the ear canal and auricle echoes will be investigated to improve the authentication system of the proposed method. Moreover, we will investigate external factors, changes over time, replay attack resistance, etc. In addition, this study was limited to an investigation using in-ear devices, and the availability of different types of devices (open-ear [2] and bone-conduction types) is also an issue for further investigation. Moreover, we need to test that our system can be combined with an inward-facing microphone.

# 7 CONCLUSION

In this paper, we proposed a system to authenticate the wearer of hearables by acquiring ear canal echo and auricle echo from the sound leakage signal of the device. As a preliminary study, we investigated the characteristics of sound leakage signals using models of the auricle and ear canal. Subsequently, we measured data from 16 participants and confirmed that the BAC scores, which indicate personal authentication performance, were in the range of 87.0%–96.7% in several scenarios.

## ACKNOWLEDGMENTS

This work is supported by JSPS, KAKENHI Grant Number 23KJ1884 and JST, PRESTO Grant Number JPMJPR2138/JPMJPR2134.

## REFERENCES

[1] Takayuki Arakawa, Takafumi Koshinaka, Shohei Yano, Hideki Irisawa, Ryoji Miyahara, and Hitoshi Imaoka. 2016. Fast and accurate personal authentication using ear acoustics. In *Proceedings of the 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA '16)*. IEEE, 1–4. https://doi.org/10.1109/APSIPA.2016.7820886

[2] BOSE. (2022.08.29. accessed). WHAT ARE OPEN-EAR HEADPHONES? https://www.bose.com/en_us/better_with_bose/open-ear-headphones.html.

[3] Yetong Cao, Chao Cai, Fan Li, Zhe Chen, and Jun Luo. 2023. HeartPrint: Passive Heart Sounds Authentication Exploiting In-Ear Microphones. In *Proceedings of the IEEE INFOCOM 2023 - IEEE Conference on Computer Communications (INFOCOM '23)*. IEEE, 10 pages.

[4] Yetong Cao, Cai Chao, Anbo Yu, Fan Li, and Jun Luo. 2023. EarACE: Empowering Versatile Acoustic Sensing via Earable Active Noise Cancellation Platform. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3, Article 47 (2023), 23 pages. https://doi.org/10.1145/3596242

[5] Seokmin Choi, Yang Gao, Yincheng Jin, Se jun Kim, Jiyang Li, Wenyao Xu, and Zhanpeng Jin. 2022. PPGface: Like What You Are Watching? Earphones Can "Feel" Your Facial Expressions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2, Article 48 (2022), 32 pages. https://doi.org/10.1145/3534597

[6] Romit Roy Choudhury. 2021. Earable Computing: A New Area to Think About. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications* (Virtual, United Kingdom) *(HotMobile '21)*. Association for Computing Machinery, 147–153. https://doi.org/10.1145/3446382.3450216

[7] Andrea Ferlini, Dong Ma, Robert Harle, and Cecilia Mascolo. 2021. EarGate: Gait-Based User Identification with in-Ear Microphones. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking (MobiCom '21)*. Association for Computing Machinery, 337–349. https://doi.org/10.1145/3447993.3483240

[8] Andrea Ferlini, Alessandro Montanari, Chulhong Min, Hongwei Li, Ugo Sassi, and Fahim Kawsar. 2022. In-Ear PPG for Vital Signs. *IEEE Pervasive Computing* 21, 1 (2022), 65–74. https://doi.org/10.1109/MPRV.2021.3121171

[9] Yang Gao, Yincheng Jin, Jagmohan Chauhan, Seokmin Choi, Jiyang Li, and Zhanpeng Jin. 2021. Voice In Ear: Spoofing-Resistant and Passphrase-Independent Body Sound Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1, Article 12 (mar 2021), 25 pages. https://doi.org/10.1145/3448113

[10] Yang Gao, Wei Wang, Vir V. Phoha, Wei Sun, and Zhanpeng Jin. 2019. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3, Article 81 (2019), 24 pages. https://doi.org/10.1145/3351239

[11] David Da He, Eric S. Winokur, and Charles G. Sodini. 2015. An Ear-Worn Vital Signs Monitor. *IEEE Transactions on Biomedical Engineering* 62, 11 (2015), 2547–2552. https://doi.org/10.1109/TBME.2015.2459061

[12] Christian Holz, Senaka Buthpitiya, and Marius Knaust. 2015. Bodyprint: Biometric User Identification on Mobile Devices Using the Capacitive Touchscreen to Scan Body Parts. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, 3011–3014. https://doi.org/10.1145/2702123.2702518

[13] David J. Hurley, Banafshe Arbab-Zavar, and Mark Nixon. 2008. The Ear as a Biometric. In *Handbook of Biometrics*. 131–150. https://doi.org/10.1007/978-0-387-71041-9_7

[14] Shunji Itani, Shunsuke Kita, and Yoshinobu Kajikawa. 2022. Multimodal Personal Ear Authentication Using Acoustic Ear Feature for Smartphone Security. *IEEE Transactions on Consumer Electronics* 68, 1 (2022), 77–84. https://doi.org/10.1109/TCE.2021.3137474

[15] Rui Liu, Cory Cornelius, Reza Rawassizadeh, Ronald Peterson, and David Kotz. 2018. Vocal Resonance: Using Internal Body Voice for Wearable Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1, Article 19 (2018), 23 pages. https://doi.org/10.1145/3191751

[16] John Middlebrooks and David Green. 1991. Sound Localization by Human Listeners. *Annual Review of Psychology* 42 (1991), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031

[17] Joseph Plazak and Marta Kersten-Oertel. 2018. A Survey on the Affordances of "Hearables". *Inventions* 3, 3 (2018). https://doi.org/10.3390/inventions3030048

[18] Alexey Sulavko. 2022. Biometric-Based Key Generation and User Authentication Using Acoustic Characteristics of the Outer Ear and a Network of Correlation Neurons. *Sensors* 22, 23 (2022). https://doi.org/10.3390/s22239551

[19] Zi Wang, Yili Ren, Yingying Chen, and Jie Yang. 2022. ToothSonic: Earable Authentication via Acoustic Toothprint. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2, Article 78 (2022), 24 pages. https://doi.org/10.1145/3534606

[20] Zi Wang, Sheng Tan, Linghan Zhang, Yili Ren, Zhi Wang, and Jie Yang. 2021. EarDynamic: An Ear Canal Deformation Based Continuous User Authentication Using In-Ear Wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1, Article 39 (2021), 27 pages. https://doi.org/10.1145/3448098

[21] World Health Organization. 2019. Safe Listening Devices and Systems.

[22] Yadong Xie, Fan Li, Yue Wu, Huijie Chen, Zhiyuan Zhao, and Yu Wang. 2022. TeethPass: Dental Occlusion-Based User Authentication via In-Ear Acoustic Sensing. In *Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications (INFOCOM '22)*. IEEE, 1789–1798. https://doi.org/10.1109/INFOCOM48880.2022.9796951