A Survey of Earable Technology: Trends, Tools, and the Road Ahead

CHANGSHUO HU†, Singapore Management University, Singapore
QIANG YANG†, University of Cambridge, United Kingdom
YANG LIU†, Florida State University, United States
TOBIAS RÖDDIGER, Karlsruhe Institute of Technology, Germany
KAYLA-JADE BUTKOW, University of Cambridge, United Kingdom
MATHIAS CILIBERTO, University of Cambridge, United Kingdom
ADAM LUKE PULLIN, University of Cambridge, United Kingdom
JAKE STUCHBURY-WASS, University of Cambridge, United Kingdom
MAHBUB HASSAN, University of New South Wales, Australia
CECILIA MASCOLO, University of Cambridge, United Kingdom
DONG MA*, University of Cambridge, United Kingdom

Earable devices, wearables positioned in or around the ear, are undergoing a rapid transformation from audio-centric accessories into multifunctional systems for interaction, contextual awareness, and health monitoring. This evolution is driven by commercial trends emphasizing sensor integration and by a surge of academic interest exploring novel sensing capabilities. Building on the foundation established by earlier surveys, this work presents a timely and comprehensive review of earable research published since 2022. We attempt to answer three core questions: (1) how has earable research evolved in recent years, (2) what enabling resources are now available, and (3) what opportunities remain for future exploration. For question (1), we analyze over one hundred recent studies to characterize this shifting research landscape, by identifying emerging sensing principles / applications, and efforts to improve accuracy and reliability over prior works. For question (2), we summarize new public datasets and hardware platforms that can be leveraged to facilitate future earable research. Finally, for question (3), we discuss open challenges and propose future directions for the next phase of earable research.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **General and reference → Surveys and overviews**.

Additional Key Words and Phrases: Earables, Sensing, Health Monitoring, Interactions, Activity, User Authentication

*Corresponding author.
†Co-primary authors with equal contribution in alphabetical order of the first names.

Authors' Contact Information: Changshuo Hu, Singapore Management University, Singapore, Singapore, cs.hu.2023@phdcs.smu.edu.sg; Qiang Yang, University of Cambridge, Cambridge, United Kingdom, qy258@cam.ac.uk; Yang Liu, Florida State University, Tallahassee, Florida, United States, yl25r@fsu.edu; Tobias Röddiger, Karlsruhe Institute of Technology, Karlsruhe, Germany, tobias.roeddiger@kit.edu; Kayla-Jade Butkow, University of Cambridge, Cambridge, United Kingdom, kjb85@cam.ac.uk; Mathias Ciliberto, University of Cambridge, Cambridge, United Kingdom, mc2514@cam.ac.uk; Adam Luke Pullin, University of Cambridge, Cambridge, United Kingdom, alp@cam.ac.uk; Jake Stuchbury-Wass, University of Cambridge, Cambridge, United Kingdom, js2372@cam.ac.uk; Mahbub Hassan, University of New South Wales, Sydney, NSW, Australia, mahbub.hassan@unsw.edu.au; Cecilia Mascolo, University of Cambridge, Cambridge, United Kingdom, cm542@cam.ac.uk; Dong Ma (corresponding author), University of Cambridge, Cambridge, United Kingdom, dm878@cam.ac.uk.

This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 2474-9567/2026/6-ART44
https://doi.org/10.1145/3810217

ACM Reference Format:
Changshuo Hu†, Qiang Yang†, Yang Liu†, Tobias Röddiger, Kayla-Jade Butkow, Mathias Ciliberto, Adam Luke Pullin, Jake Stuchbury-Wass, Mahbub Hassan, Cecilia Mascolo, and Dong Ma*. 2026. A Survey of Earable Technology: Trends, Tools, and the Road Ahead. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 2, Article 44 (June 2026), 41 pages. https://doi.org/10.1145/3810217

1 INTRODUCTION

Earables are wearable computing devices worn in or around the ear, capable of supporting audio playback, user interaction, and physiological sensing. Among various earable form factors, true wireless stereo (TWS) earbuds have become the most commercially successful and widely adopted platform [143]. The global market for TWS earbuds has grown exponentially and is expected to continue accelerating in the coming years. Forecasts from The Business Research Company project that the market will expand from 89.6 billion USD in 2024 to 121.91 billion USD in 2025, and further reach 415.69 billion USD by 2029, with a compound annual growth rate of 35.9% [6]. This rapid growth is driven not only by the widespread adoption of earphones but also by ongoing technological advances from the academic community, the integration of diverse sensors, and a growing awareness of health and fitness [24, 117]. These trends reflect a fundamental shift: *earbuds are evolving from basic audio playback devices into multifunctional, intelligent systems that support entertainment, communication, and personal wellness.*

To better understand how commercial priorities for earbuds have evolved, we analyzed promotional materials for 278 TWS models released between 2020 and 2025 by major electronics companies (e.g., Apple, Samsung, Sony, Huawei) as well as specialized audio brands (e.g., Bose, Beats, JBL, Sennheiser). For each earbud model, we extracted 4–8 keywords from the promotional content with a focus on product features. We then identified and counted the number of keywords related to contextual and human-centric sensing[1]. Figure 1 illustrates the annual trend in the prevalence of these sensing-related keywords. We observe a clear upward trajectory, indicating that manufacturers are increasingly emphasizing sensing and context-awareness features as key differentiators in the TWS market. This shift highlights a broader transition from audio-only devices toward multi-functional, sensor-rich wearables that support diverse applications in health, fitness, and ambient intelligence.

This transformation is grounded in several intrinsic advantages that make earbuds an ideal platform for both user interaction, behavioral sensing, and physiological monitoring. First, as devices primarily used for audio playback and telephony, *earbuds are already familiar to users*, who naturally interact with them through voice commands or touch controls [62, 128]. Second, *the usage scenarios for earbuds have become increasingly diverse* [147]. They are now commonly worn during commuting for noise cancellation, during exercise for music streaming, and in work settings for meetings or recording. Recent studies also report that users in their 20s to 40s typically wear earbuds nearly every day [141] and for about 50 to 60 minutes per day [7]. These everyday contexts provide valuable opportunities to monitor and analyze user activities in real-world environments. Third, *earbuds occupy a uniquely advantageous position on the body.* Situated on the head and in close proximity to facial muscles, cranial bones, the auditory canal, and major physiological structures such as the brain and large blood vessels, they can access a variety of internal signals [143]. This feature distinguishes earables from wrist- [27] or finger-worn wearables [175] by not only measuring peripheral signals such as pulse or skin temperature but also deeper physiological cues like respiration acoustics, bone-conducted speech, heart sounds, and potentially brain-related activity. Fourth, *their relatively stable placement on the human head helps reduce motion artifacts* that are commonly encountered at the wrist, thereby improving signal quality and sensing robustness [70]. Finally, *the dual-element design (i.e., left and right earbuds) increases the information gain of sensing signals* by providing multi-stream and spatially diverse inputs [25, 77].

[1]These sensing-related keywords include: Adaptive Sound Adjustment, Environmental Awareness, Fitness Tracking, Hearing Health, Heart Rate Monitoring, Sleep Monitoring, SpO_2 Detection, Spatial Audio, Speech Enhancement, Stress Monitoring, Temperature Monitoring, Touch Control, and Voice Assistant.

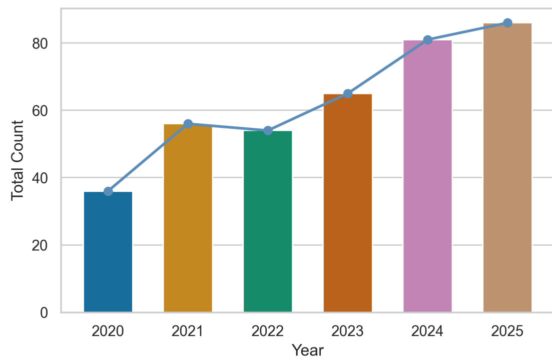

Fig. 1. Annual trend of sensing-related keywords mentioned in earable product descriptions from 2020 to 2025.

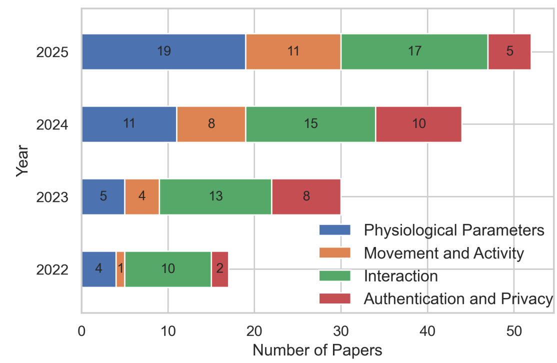

Fig. 2. Total number of papers by categories published on earable technologies from 2022 to 2025.

While commercial development has largely begun with basic earable sensing functions, such as gesture control [63, 128] and heart rate monitoring [2], in consideration of reliability and user adoption, the academic community has been more proactive in exploring the broader sensing potential of earbuds, particularly over the past three years. We acknowledge that Röddiger *et al.* [143] provided a comprehensive survey of this research field up to and including 2021, offering a taxonomy of phenomena and sensing modalities enabled by earables across various interdisciplinary communities. This survey highlighted that earables can sense a wide range of physiological signals, including motion [77, 89, 152], respiration [76, 117], heartbeat [24, 25], body-conducted sounds [178, 186], as well as subtler cues such as food intake [59, 97], brain activity [18].

Building on this foundation, the pace of academic progress has accelerated dramatically since then—possibly spurred by the impact of this very survey. To our surprise, more than one hundred new papers on earable sensing have been published since 2022 in major ubiquitous computing venues[2] with an upward annual trend (as shown in Figure 2), reflecting a rapid expansion of both interest and technical capabilities within the community. *Given the scale and speed of these developments, we believe there is a clear and timely need for an updated survey to address three fundamental questions relevant to both existing researchers and newcomers to this field*:

- What kind of earable research has emerged since 2022, and how does it differ from prior work? (*Trend*)
- What resources (e.g., hardware platforms and datasets) are currently available to facilitate future earable research? (*Tools*)
- Is earable computing a saturated research domain, or does it remain a promising frontier? If the latter, what are the key directions for the next phase of earable research? (*Road ahead*)

To answer these questions, in this survey, we conduct a comprehensive review of over one hundred recent papers on earable sensing published between 2022 and 2025. Overall, our findings highlight that, over the past three years, researchers have discovered a variety of novel sensing principles, introduced numerous new earable sensing applications, improved the performance of existing sensing tasks, and developed significant new resources to further advance research in the field. Concretely, for applications previously examined in earlier surveys, we revisit each category and analyze how recent works advance prior efforts, highlighting improvements such as enhanced accuracy, increased robustness, finer-grained contextual inference, and lightweight implementations that reduce delay and energy overhead. For emerging applications not covered before, we examine their motivations, underlying sensing principles, and the types of sensors employed, offering a structured account of how these innovations expand the functional scope of earables. In addition, we synthesize enabling resources such as

[2]Including IMWUT, CHI, MobiCom, MobiSys, SenSys, PerCom, INFOCOM, UIST, ISWC, and associated workshops.

hardware platforms and public datasets, and conclude with a discussion of open challenges and promising future directions for the next phase of earable research.

The rest of the paper is organized as follows: Section 2 introduces our methodology for paper retrieval and selection. Section 3 presents a taxonomy of earable sensing applications. Sections 4 to 7 provide an in-depth review of recent works across four major domains: physiological parameters and health (Section 4), movement and activity (Section 5), interaction (Section 6), and authentication and privacy (Section 7). Section 8 discusses enabling technologies and public resources such as hardware and datasets. Finally, Section 10 outlines future research directions, before concluding the paper in Section 11.

2 METHODOLOGY

This survey aims to provide a comprehensive overview of recent research efforts on earable sensing, with a particular focus on works published between 2022 and 2025 in major ubiquitous computing and wearable systems venues. We adopt a structured literature review approach that combines keyword-based retrieval, multi-stage screening, backward citation tracing, and thematic analysis. The final literature search was conducted on January 20, 2026. Only peer-reviewed publications were considered in this survey. This section outlines the steps taken to select and analyze the relevant literature.

2.1 Paper Retrieval

We retrieved research articles from major digital libraries and publishers, including the ACM Digital Library (ACM-DL), IEEE Xplore (IEEE-X), as well as journals and conference proceedings published by Nature, Science, Springer, and MDPI. Together, these sources host a substantial portion of the literature in ubiquitous computing, human-computer interaction, and wearable sensing.

Our search targeted paper titles, abstracts, and metadata using an extensive list of keywords related to earable sensing. These included general terms such as *earable*, *hearable(s)*, *ear-worn*, *ear(-)mounted*, *ear(-)attached*, *ear(-)based*, *in-ear device*, *earbud(s)*, *earphone(s)*, *earpiece(s)*, *headphone(s)*, *microphone(s)*, *IMU(s)*, *PPG(s)*, and *photoplethysmography(s)*. To ensure broader coverage across application domains, we also constructed compound queries by combining ear-related terms with task-specific keywords (e.g., *ear-based AND authentication*, *ear-based AND gesture*).

2.2 Selection Criteria, Filtering, and Backward Chaining

The paper selection process consisted of multiple screening stages, including initial keyword-based retrieval, title and abstract screening, full-text review, and backward citation tracing.

We first applied the following inclusion criteria:

(1) the paper describes a device that is worn in, on, or around the ear; and
(2) the sensing takes place at or near the ear, either through direct contact or close-range interaction.

We then excluded works that:

(1) focus solely on voice and audio interfaces (e.g., active noise cancellation) or on the technical design of audio hardware without a sensing objective;
(2) are not peer-reviewed publications, such as preprints (e.g., arXiv), patents, dissertations, or technical reports;
(3) involve larger head-worn systems (e.g., virtual or augmented reality headsets) that do not conform to typical earable form factors;
(4) are not written in English; or
(5) do not include a functioning prototype or system-level evaluation.

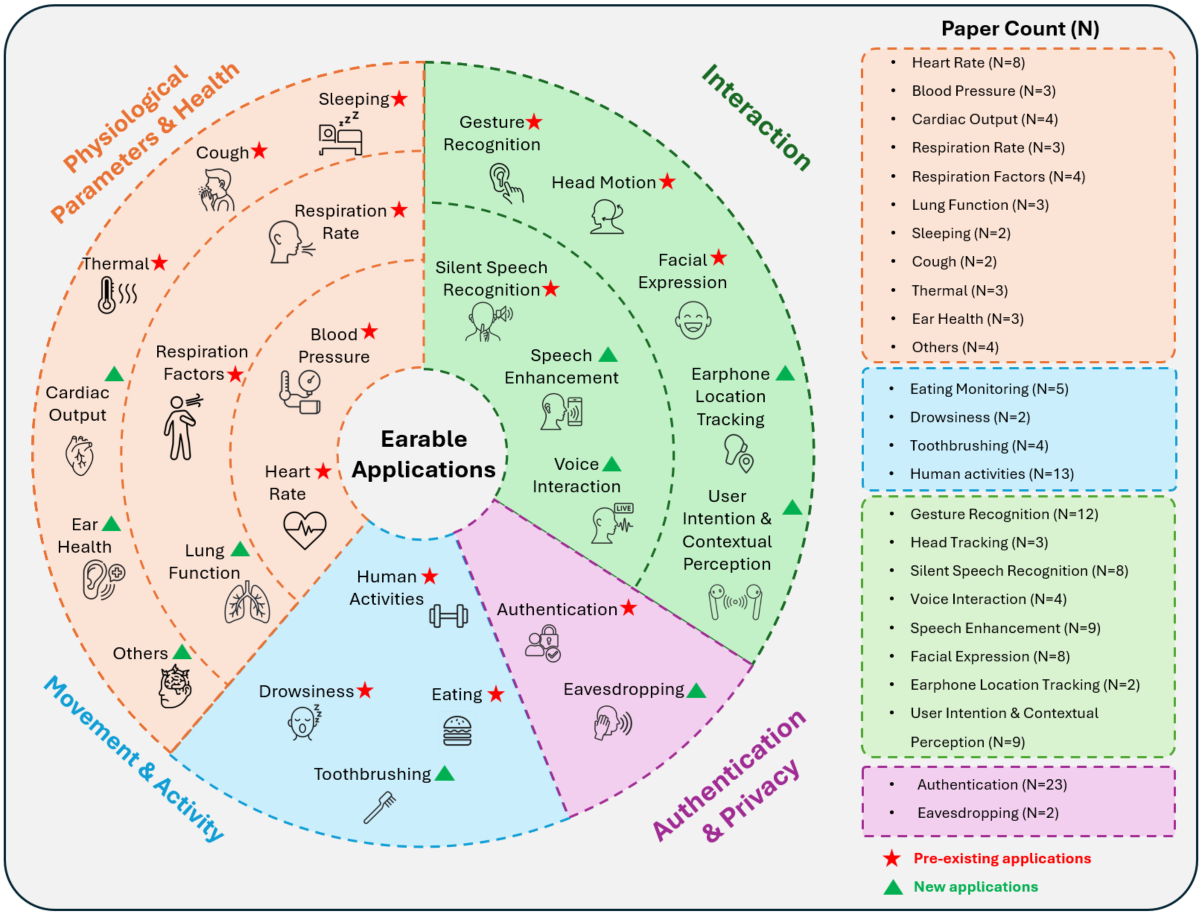

Fig. 3. Paper overview. The map is organised by phenomena that can be captured with earables. Phenomena are clustered based on shared sensing themes or physiological systems, and grouped into four main categories: Physiological Parameters and Health (Section 4), Movement and Activity (Section 5), Interaction (Section 6), and Authentication and Privacy (Section 7).

All candidate papers were independently reviewed by multiple authors based on their titles and abstracts, and full texts were examined when necessary to ensure consistent application of the inclusion and exclusion criteria. After this screening process, we conducted backward citation tracing on the reference lists of the selected papers to identify additional relevant studies that were not retrieved during the initial database search. Each selected paper was then annotated with key attributes, including publication year, venue, sensing modality, sensor type, targeted application, and evaluation methodology. These attributes formed the basis for our comparative analysis, enabling us to categorize application domains, identify emerging research trends, and assess both the technical depth and practical relevance of prior work. In total, following this multi-stage selection and screening process, we identified 153 peer-reviewed papers, which collectively form the basis of the analyses presented in this survey.

3 EARABLE SENSING TAXONOMY

We follow a taxonomy structure similar to the survey by Röddiger et al. [143] on earable sensing and categorize applications into four primary domains: Physiological Parameters and Health, Movement and Activity, Interaction,

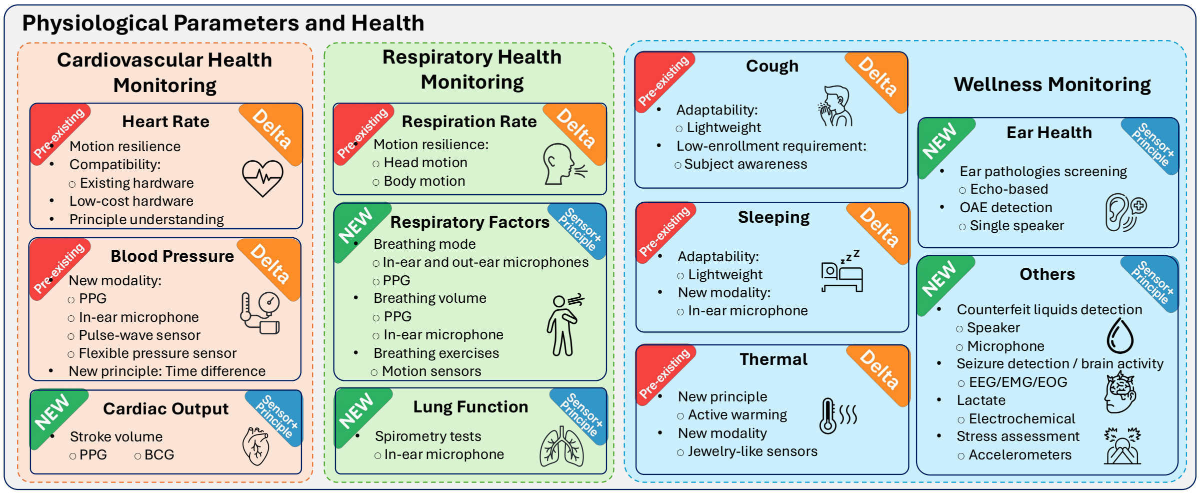

Fig. 4. Summary of recent advancements in earable-based physiological sensing. We categorize works into (1) **Pre-existing tasks** explored before 2021, where post-2021 studies introduce incremental **Delta** improvements; and (2) **Emerging applications**, marked as **NEW**, representing novel sensing opportunities uniquely enabled by the earable form factor after 2021.

and Authentication and Privacy, as illustrated in Figure 3. Each domain organizes sensing applications based on the human phenomena being monitored or enabled by earables.

To capture how the field is evolving, we further divide applications into two classes based on whether they appeared in previous earable survey [143]. The first class includes **pre-existing applications** (labeled with *red pentagram* in Figure 3) that had been explored prior to 2022, such as gesture recognition or heart rate monitoring. For these, we analyze how recent works have addressed past limitations (such as sensitivity to noise and motion, high enrollment costs, energy inefficiency, and etc.) and identify improvements in accuracy, latency, or robustness. The second class highlights **new applications** that have emerged since the last major survey [143] (labeled with *green triangles* in Figure 3), such as ear health detection and speech enhancement. Here, we focus on aspects such as motivation behind the application, their novel sensing principles, and unique hardware configurations.

After reviewing advancements across each domain, we conclude with a brief ***Remarks*** section that distills key technical trends of the evolving landscape of earable sensing. These remarks highlight a clear shift toward passive sensing, energy efficiency, and continual learning, along with growing emphasis on usability, cross-user generalization, and multimodal fusion.

4 PHYSIOLOGICAL PARAMETERS AND HEALTH

This section reviews recent progress in physiological parameters and health monitoring with earables (Figure 4).

4.1 Cardiovascular Health Monitoring

Continuous cardiovascular health monitoring is essential not only in clinical settings, but also for daily-life or at-home health management [149]. By regularly tracking indicators such as heart rate, heart rate variability, and blood pressure, both healthcare professionals and individuals can promptly detect changes in cardiac function and take timely action [129]. Moreover, continuous cardiovascular monitoring in everyday life can enhance physical activity levels and overall cardiovascular health [137].

Heart Rate Metrics. Recent earable-based cardiovascular research has primarily focused on heart rate metrics: heart rate (HR), heart rate variability (HRV), and cardiac waveforms, with growing attention to three directions: (i) leveraging low-cost, widely available earphones [37], (ii) achieving robustness under real-world motion and noise [23, 24, 51, 118, 184, 206], and (iii) deepening physiological understanding [43, 176] as summarized in the supplementary material. Fan *et al.* [51] introduced *APG*, which uses ultrasonic probes from active-noise-cancelling (ANC) headphones to sense heart activity through built-in microphones. Building on low-cost hardware reuse, Chen *et al.* [37]'s *Asclepius* system turns ordinary wired earphones into electronic stethoscopes using an inexpensive plug-in board and signal-processing pipelines to extract cardiac sounds and HR/HRV. Focusing on motion robustness, Butkow *et al.* [23, 24] developed *hEARt*, a deep-learning framework that enhances in-ear microphone signals for HR estimation under dynamic conditions. Furthermore, Zhu *et al.* [206] proposed *VitalEar*, a deep-learning-based system that simultaneously estimates heart and respiratory rates during aerobic exercise using in-ear microphones. Complementarily, Liu *et al.* [118] introduced *EarFusion*, a quality-aware fusion framework that combines in-ear acoustic signals with PPG to improve heart rate estimation under motion and noise. In a complementary direction that also emphasizes form factor and wearability, Xue *et al.* introduced *PPG Earring* [184], a wireless smart earring that integrates an earlobe PPG sensor into a jewelry-like design to enable comfortable, continuous heart monitoring while mitigating the motion-artifact limitations of wrist/ring PPG during daily activities and exercise. On the physiological modeling front, Christofferson *et al.* [43] provided the first systematic analysis of how cardiac vibrations propagate to the ear canal, revealing insights that inform sensor placement and algorithm design, while Waters *et al.* [176] developed a deep-learning model that segments in-ear cardiac sounds into heartbeat phases, demonstrating the feasibility of accurate heart sound analysis using earphone microphones. *Together, these studies highlight the field's evolution toward low-cost hardware reuse, motion-resilient algorithms, and physiologically grounded earable-based cardiovascular monitoring.*

Blood Pressure. Parallel research has extended earable sensing beyond heart rate by exploring blood pressure (BP) estimation [19, 45, 102, 111, 194, 204] as summarized in the supplementary material. Balaji *et al.* [19] proposed *Stereo-BP*, which estimates BP from pulse arrival time differences between PPG signals captured at both ears. Moreover, Zhao *et al.* [204] introduced *HearBP*, which leverages bone-conducted heart sounds recorded by in-ear microphones and deep learning–based denoising for cuffless BP estimation. Moving beyond optical/acoustic sensing, Kim *et al.* [102] developed an earphone-form-factor system that integrates a self-powered triboelectric pulse-wave sensor for continuous BP monitoring. Similarly, Li *et al.* [111] presented *FlexBP*, which enables continuous BP monitoring using a single-point, skin-conformal in-ear flexible pressure sensor to capture arterial pulsations for data-driven SBP/DBP estimation. Complementing direct BP regression, Christofferson *et al.* [45] proposed *ArtEARial*, which reconstructs the arterial pressure waveform from in-ear cardiac audio recorded by an ANC earbud microphone, enabling pulse-wave-analysis features that can support downstream BP assessment. Additionally, Yi *et al.* [194] showed that ordinary consumer earphones/headphones can be repurposed to capture low-frequency pulse pressure waves in the ear canal via simple signal separation, enabling access to peak-to-peak intervals and pulse-waveform features that could support downstream BP and pulse-wave-analysis applications. *Together, these studies demonstrate the feasibility of in-ear, non-invasive BP monitoring, highlighting the shift toward comfortable, continuous, and clinically meaningful cardiovascular assessment.*

Cardiac Output. As a new and emerging application, recent studies [205] have extended earable sensing to estimate cardiac output (CO), a key indicator of the heart's pumping efficiency and overall cardiovascular performance. Zhou *et al.* [205] proposed *EarCO*, the first multimodal earbud-based system for non-invasive CO monitoring. By jointly analyzing PPG and BCG signals captured from commodity earbuds, EarCO employs a feature fusion and guidance framework that integrates physiological biomarkers with deep representations to enhance interpretability and accuracy. *This emerging direction broadens earable sensing from tracking cardiac rhythms to assessing comprehensive cardiac function using earphones.*

Remarks. Overall, recent work demonstrates that commodity earphones can: (1) accurately estimate HR/HRV, BP, and even CO using in-ear acoustic, optical, or motion signals; (2) maintain robustness under motion and ambient noise; and (3) leverage low-cost, physiologically grounded designs for practical, continuous cardiovascular monitoring in daily life.

4.2 Respiratory Health Monitoring

Continuous respiratory health monitoring is essential not only in clinical settings, but also for everyday scenarios like sports training, sleep tracking, and overall wellness [125, 165]. By closely observing indicators such as breathing rate, airflow patterns, and other aspects of lung function, both healthcare professionals and individuals can detect subtle changes early and respond promptly [165]. Recent research in earable-based respiratory health monitoring has primarily focused on several key physiological metrics, including respiration rate (RR) [10, 117, 146, 206], breathing modes [76, 146], tidal volume [76], breathing volume [119], and lung function assessment [30, 177, 179] as summarized in the supplementary material.

Respiration Rate. Recent studies [10, 117, 206] have advanced earable-based respiration monitoring toward motion-rich, real-world scenarios. Ahmed *et al.* [10] proposed a multimodal system that combines motion and acoustic sensing on commercial earbuds to maintain reliable respiration tracking during head movements. Taking a step further, Liu *et al.* [117] introduced *RespEar*, which leverages in-ear microphones and physiological couplings to enable robust respiration monitoring across both sedentary and active conditions. Most recently, Zhu *et al.* [206] proposed *VitalEar*, a deep-learning-based system that simultaneously estimates heart and respiratory rates during aerobic exercise using in-ear microphones. *Together, these efforts demonstrate the feasibility of accurate, motion-resilient respiratory sensing using earphones.*

New Respiratory Factors. Beyond respiration rate, recent work [76, 119, 140, 146] has explored richer aspects of respiratory health. Hu *et al.* [76] introduced *BreathPro*, which recognizes nasal and oral breathing modes during running using in- and out-ear microphones. Meanwhile, Romero *et al.* [146] proposed *OptiBreathe*, which leverages in-ear PPG signals to monitor respiratory rate, breathing phases, and tidal volume. Additionally, Liu *et al.* [119] proposed *EarMeter*, which uses in-ear microphones and a learning-based pipeline to continuously estimate respiration volume during natural breathing across varying intensities and daily activities. Recently, Rahman *et al.* [140] developed *MindfulBuddy*, an earbud motion–based system that extracts comprehensive breathing biomarkers, including depth, symmetry, and breath-holding, for real-time biofeedback during breathing exercises. *Together, these studies expand earable sensing toward comprehensive respiratory health and breathing assessment.*

Lung Function. Recent studies [30, 177, 179] have also explored the potential of earables for lung function assessment. Xie *et al.* [177] introduced *EarSpiro*, the first earphone-based system that reconstructs full flow–volume curves from in-ear acoustic signals during pulmonary function tests (PFTs). Taking a step further, Xu *et al.* [179] proposed *EasySpiro*, which estimates PFT indicators from sub-maximal exhalations by reconstructing ideal maximal patterns guided by IMU-based effort encoding. Meanwhile, Cao *et al.* [30] developed *ESPIRO*, which infers pulmonary function indices from normal speech captured by earphones, linking glottal flow features to lung performance without requiring forced breathing. *These works extend earable sensing from respiratory rate tracking to convenient, natural, and clinically relevant pulmonary function evaluation.*

Remarks. Collectively, recent studies demonstrate that earables can: (1) achieve accurate and motion-resilient respiration rate monitoring; (2) capture richer respiratory factors such as breathing modes, tidal volume, and mindfulness-related biomarkers using microphones, motion sensors, or PPG; and (3) leverage physiology-informed learning models to derive clinically meaningful lung function metrics from earphones.

4.3 Wellness Monitoring

Beyond cardiovascular and respiratory monitoring, recent research has leveraged earables to track a wider range of wellness-related factors, including coughing [172, 201], sleep quality [44, 65], and thermal regulation [104, 183], as well as newer applications such as ear-health assessment [31, 83, 90] and other emerging use cases [18, 86, 114, 181] as summarized in the supplementary material.

Cough. Recent work [172, 201] has leveraged earables for continuous cough monitoring to support airway health assessment in daily life. Wang *et al.* [172] introduced *HearCough*, which performs on-device cough detection using lightweight neural networks on earbud microcontrollers for privacy and energy efficiency. After that, Zhang *et al.* [201] proposed *EarCough*, which adds subject awareness by distinguishing the wearer's coughs from ambient ones using dual-microphone inputs. *Together, these systems advance low-power, personalized, and privacy-preserving cough monitoring with earphones.*

Sleeping. Earables have also been explored for sleep assessment [44, 65]. Christofferson *et al.* [44] used ANC earphones with dual microphones to classify sleep-related sounds such as snoring and movement through lightweight, on-device models. Furthermore, Han *et al.* [65] proposed *EarSleep*, which leverages in-ear microphones and a dual-stream learning framework to capture body-conducted sounds for fine-grained sleep stage detection. *Together, these systems highlight the potential of earables for unobtrusive, multi-feature, and at-home sleep analysis.*

Thermal. Complementing physiological and behavioral sensing, recent work [104, 183] has explored earables for thermal comfort and long-term temperature monitoring. Knierim *et al.* [104] introduced *Warmth on Demand*, a thermoactive headphone that locally warms or cools the ear region to enhance comfort in varying environments. Meanwhile, Xue *et al.* [183] developed *Thermal Earring*, a jewelry-like wireless sensor that continuously tracks earlobe temperature for weeks on a single charge. *These studies highlight the potential of earables for unobtrusive, fashionable, and continuous thermal well-being monitoring.*

New Applications—Ear Health. In parallel, earables are being extended toward clinical-grade ear health assessment [31, 83, 90]. Jin *et al.* [90] introduced *EarHealth*, which detects common ear conditions such as eardrum rupture, otitis media, and earwax blockage using acoustic echoes captured by earphones. Further expanding diagnostic capabilities, Chan *et al.* [31] developed *OAEbuds*, a low-cost wireless earbud system that screens cochlear function by detecting otoacoustic emissions with a single speaker. Most recently, Huang *et al.* [83] proposed *EarCSI*, a reconstruction-driven framework that models ear canal geometry and reflections from passive audio to detect fine-grained tympanic membrane changes using commercial headphones. *Together, these works advance affordable, in-situ, and fine-grained ear health sensing using earphones.*

New Applications—Others. Finally, researchers have begun repurposing earables for broader applications beyond vital-sign and clinical monitoring [18, 86, 114, 181]. Li *et al.* [114] introduced *ASLiquid*, which uses earphone speakers and microphones to detect counterfeit liquids through acoustic resonance analysis. In another direction, Aziz *et al.* [18] developed *EarSD*, an ear-worn system that captures EEG, EMG, and EOG signals for real-time epileptic seizure detection. Additionally, Islam *et al.* [86] introduced *BallistoBud*, which uses earbud accelerometers to extract BCG signals for HRV–based stress assessment. Xu *et al.* [181] presented an in-ear integrated sensor array that combines electrophysiological sensing (e.g., EEG/EOG/EDA) with electrochemical sweat sensing to simultaneously track brain activity and lactate using a user-generic earphone-mounted form factor. *Together, these works showcase the expanding potential of earables for safety, wellness, and clinical monitoring in everyday life.*

Remarks. Collectively, recent studies demonstrate that earables can: (1) perform privacy-preserving, on-device sensing for continuous cough and sleep monitoring; (2) integrate comfort and wellness features such as thermal regulation and long-term temperature tracking; and (3) expand into new domains including ear health assessment, stress and seizure detection, and counterfeit-liquid identification.

5 MOVEMENT AND ACTIVITY

Earable devices have shown growing potential in capturing diverse aspects of human movement and activity. As shown in Figure 5, this section reviews recent progress in earable applications for movement and activity sensing.

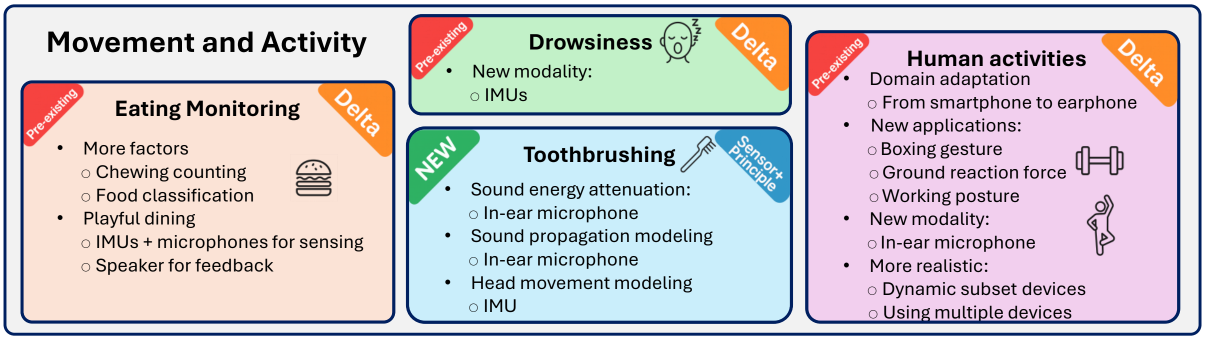

Fig. 5. Summary of recent advancements in earable-based movement and activity sensing.

5.1 Eating Monitoring

Recent work [59, 75, 97, 134, 170] has advanced earable sensing for unobtrusive eating monitoring and experience enhancement as summarized in the supplementary material. Ketmalasiri *et al.* [97] proposed *IMChew*, which uses earphone IMUs to detect and count chewing activity for personalized dietary tracking. Furthermore, Srivastava *et al.* [59] introduced *BiteSense*, which combines IMU-based motion analysis with temporal modeling to infer food type, texture, and eating pace. Complementing IMU-based approaches, Hidaka *et al.* [75] used an earphone-style light-sensor "mastication meter" placed at the outer ear to quantify chewing duration and chew counts, enabling objective analysis of masticatory behaviors in people with obesity. In another direction, Wang *et al.* [170] presented *GustosonicSense*, which fuses in-ear audio and motion sensing to provide playful, real-time sound feedback that enhances dining enjoyment and mindfulness. Additionally, Nakamura *et al.* [134] proposed *Eat2pic-Mobile*, which uses commodity wireless earphones (microphones) and a smartphone to detect chewing and provide art-based visual feedback (Chew-Draw) that encourages mindful chewing and balanced eating. *Together, these studies demonstrate the potential of earables for fine-grained, socially acceptable, and engaging dietary monitoring.*

Remarks. In summary, recent studies have advanced earable eating monitoring from simple chew detection to: (i) quantitative meal metrics such as chew count and pace, (ii) fine-grained dietary profiling that recognises food type, and (iii) experiential augmentation that enriches dining through real-time audio feedback.

5.2 Drowsiness

Fatigue and drowsiness detection are vital for safety and productivity in tasks requiring sustained attention. Brown *et al.* [22] introduced an IMU-based earable system that detects yawning through head and jaw motion, offering a privacy-preserving and low-cost alternative to vision-based approaches. Complementing motion-based cues, Kaveh *et al.* [95] developed a wireless ear-EEG platform with dry, user-generic in-ear electrodes and compact hardware for drowsiness classification, demonstrating the feasibility of population-trained models with in-ear electrophysiology. Together, these studies demonstrate the potential of earables for unobtrusive and reliable fatigue monitoring in daily activities, as summarized in the supplementary material.

5.3 Toothbrushing

Many prior studies have explored using wrist-worn IMU sensors, such as smartwatches, to track brushing activity. However, these approaches are susceptible to noise from general hand movements and non-brushing activities, limiting their accuracy in practical scenarios [143]. Recently, earables have been used to monitor toothbrushing activity in real-world scenarios due to the ubiquitous presence of sensors and the location close to the mouth.

BrushBuds [185] addressed the limitations of wrist-worn IMUs by directly using the motion sensors in earphones. By utilizing accelerometers and gyroscopes, the system captured subtle head and jaw movements during brushing, providing a more stable sensing method. The result shows BrushBuds can achieve a detection accuracy of 84.3% for six different mouth regions. While BrushBuds successfully distinguished different brushing regions, it struggled to differentiate between surfaces of the same tooth due to the similar head movements, limiting its precision. To overcome the limitations of IMU-based tracking, ToothFairy [169] repurposed earphones to collect bone-conducted sound signals for detecting toothbrushing locations at a tooth-level resolution (90.5%). By analyzing how the intensity of the vibration sound is attenuated after the propagation of different brushed teeth, the system achieved precise tracking but was limited to electric toothbrushes due to their consistent vibration patterns. Building on the similar principle of ToothFairy, SmarTeeth [186] expanded the in-ear audio-based approach to manual toothbrushing. It used variations in the way the sound traveled through the skull and reached the ear canal based on the location of the brushing, allowing SmarTeeth to detect brushing activities beyond electric toothbrushes, making it a more generalizable solution for everyday users. HearForce [187, 188] advances this direction by demonstrating, for the first time, that toothbrushing force can be estimated from in-ear audio with a pair of earbuds.

Remarks. With these foundational studies, future research can advance in several directions: (1) recognizing different brushing techniques such as up-and-down, side-to-side, or circular motions. (2) Combining acoustic sensing with IMU data, barometric pressure, and other physiological signals to improve the performance and robustness of brushing detection. (3) By integrating real-time audio feedback, gamification, and voice guidance, earable technology can make brushing more engaging and effective for children. We summarized these works in the supplementary material.

5.4 Human Activity Recognition

Recent advances in earable devices have opened new avenues for non-intrusive and continuous activity sensing. Earables are uniquely positioned on the head, allowing them to unobtrusively capture subtle body and head movements, vibrations, and even bone-conducted signals. Their integration into daily life and compatibility with off-the-shelf hardware make earables a practical platform for on-the-go activity recognition. Building on this potential, Hu *et al.* [77] proposed a system that uses the in-ear microphone to detect different foot strike patterns during running. A preliminary study [92] validates an earbud for gait analysis, showing excellent agreement with motion-capture and force-plate references across key walking and running parameters, and demonstrating its feasibility as a low-cost, portable alternative to conventional lab-based systems. WalkEar [158] further proposed a holistic system for spatio-temporal, kinetic, and asymmetry gait parameter monitoring with earable IMUs. Similarly, Motion2Press [142] introduces a cross-modal learning framework that reconstructs plantar pressure, ground reaction force (GRF), and center of pressure (COP) from IMU signals, enabling low-cost and real-world gait analysis. GCCRR [182] proposes a method for segmenting gait cycles from short IMU sequences using ear-worn sensors to support home-based rehabilitation. Bian *et al.* [21] designed a low-power step-counting solution utilizing the body area electric field acquired by a novel electrostatic sensing unit. For head gesture recognition, Sepanosian and Incel [152] introduced a real-time boxing gesture recognition system using IMU sensors on OpenEarable. Meanwhile, Moschina *et al.*[133] explored earable accelerometers for vertical jump testing, achieving a mean absolute error of only 0.04m in estimating jump height.

For more comprehensive motion tracking, Ear2Pos [166] estimates full-body pose using only two IMUs integrated into earbuds. IMUPoser [130] estimates full-body pose using any available subset of consumer devices (earbuds, phones, watches). Building on this, MobilePoser [180] integrates pose estimation and global translation using a hybrid deep learning and physics-based pipeline. It achieves real-time performance on smartphones and demonstrates robustness with as few as one earable device. Moving beyond individual actions, EarDA [120] proposes an adversarial domain adaptation framework to improve HAR accuracy and data efficiency on earables by transferring knowledge from smartphone datasets. It combines Bi-LSTM feature extraction with filtering techniques to mitigate noise from unpredictable head motion, achieving nearly 89% accuracy. Choi *et al.* [39] develops and validates an earphone-based human activity recognition system that automatically measures

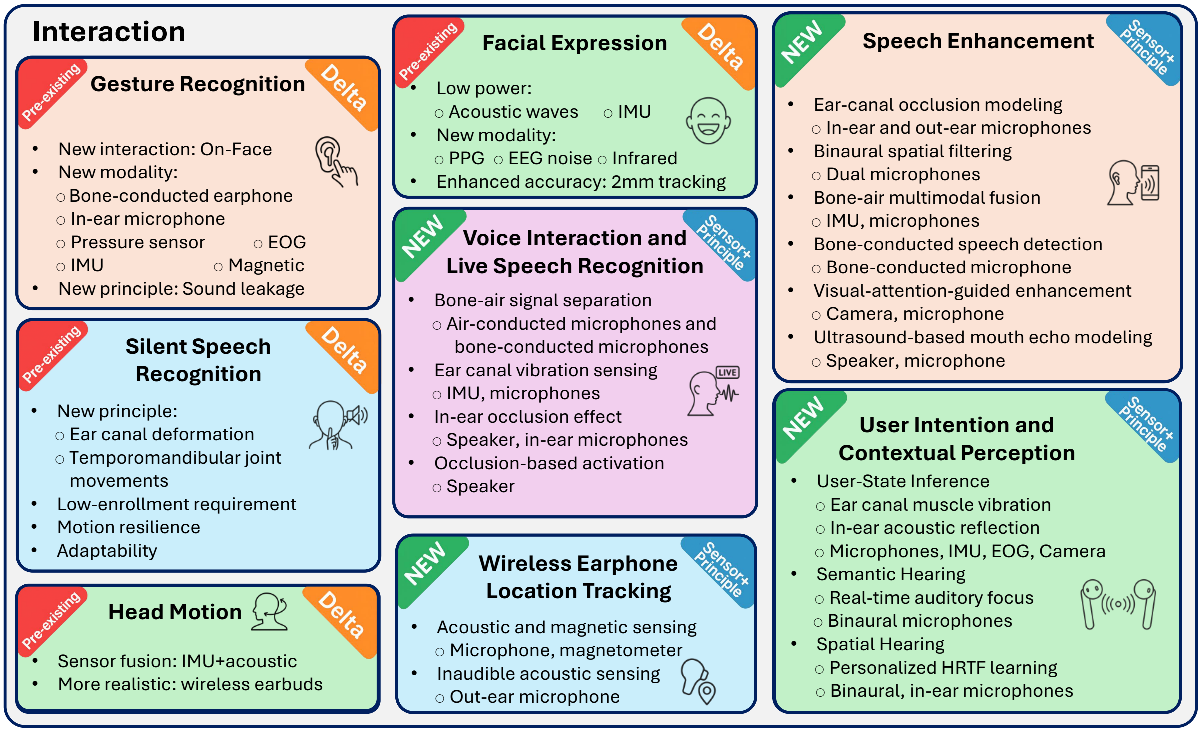

Fig. 6. Summary of recent advancements in earable-based interaction.

short physical performance battery (SPPB) components: balance, gait speed, and chair-stand, showing high agreement with manual assessments and enabling accessible, remote physical-function evaluation for older adults. Complementing this focus on physical function assessment, EPICS [8] introduces a dual-sensor earable–waist system to recognize deskwork postures in real time and deliver corrective feedback, achieving high accuracy and demonstrating strong potential for promoting healthier ergonomic behavior.

Remarks. These studies collectively highlight the growing trend of earable-based HAR. However, challenges remain in dealing with sensor variability, movement noise, and limited labeled data. Future research should focus on multi-modal fusion, domain adaptation across users and devices, and large-scale deployment studies. In particular, developing standardized benchmarks and exploring on-device learning approaches will be key to bringing earable activity recognition into real-world applications at scale. We summarized these works in the supplementary material.

6 INTERACTION

This section summarizes recent advances in earable interaction modalities, which span both refinements of existing interaction paradigms, such as gesture recognition, facial expression, and head motion, and the emergence of new capabilities like speech enhancement, voice interaction, and earphone tracking (Figure 6).

6.1 Gesture Recognition

Gesture recognition offers a more intuitive and socially acceptable alternative to traditional touch or voice interfaces. Early work in this space largely relied on IMU-based gesture recognition, which remains a foundational direction. EarBender [14] and Uni-Ear [202] leveraged commercial earables' built-in IMUs to detect natural hand-to-ear gestures such as taps and swipes. Building on this, Sato *et al.* [150] emphasized user-driven personalization, conducting a Gesture Elicitation Study to collect user-defined gestures and designing an IMU-based recognition model that achieved over 91% accuracy while accounting for personalized gestures and interaction-area constraints. Pushing further toward user-independent performance, Ui-Ear [202] introduced a novel on-face gesture recognition method based on vibration sensing. By combining IMU signals with domain-adversarial learning, they achieved strong cross-user generalization with 82.3% accuracy.

To expand the sensing capability beyond IMUs, MAF [191], LeakyFeeder [190], and Handleap [72] employed active acoustic sensing using earphones, capturing perturbations in the emitted acoustic field caused by hand gestures. Similarly, EarHover [160] proposed a mid-air gesture recognition method based on Doppler shifts in leaked audio. Unlike prior work focused on contact gestures, it achieved 88.8% accuracy in real-world scenarios while remaining compatible with audio playback. For in-ear sensing, OESense [123] introduced a powerful method that exploits the occlusion effect inside the ear canal. By capturing low-frequency vibrations using inward-facing microphones, it robustly recognized subtle face tapping gestures with up to 97% recall, even in noisy environments. PalateTouch [203] leverages in-ear acoustic sensing technology to detect gestures resulting from the interaction between the tongue and the palate. It can capture subtle ear canal deformation and recognize various palate gestures used for interaction. Complementing this, BudsID [100] uses an integrated magnetometer to sense the change in the magnetic field caused by a magnetic ring worn on a user's middle finger and to recognize different tapping fingers (among index, middle, and ring). EarEOG [116] uses around-ear electrooculogram (EOG) signals to detect facial gestures and control the wheelchair. Alternatively, Iguma *et al.* [85] innovated with the use of atmospheric pressure changes within the ear canal, enabling detection of both touch and non-touch interactions (*e.g.*, ear covering) with up to 99% accuracy. From a design perspective, Panda *et al.* [135] expanded the interaction space by treating headphones as rich, sensor-enhanced surfaces. Their design exploration incorporated multi-modal sensing (capacitive touch, depth sensing, head orientation), showing how such integration can support complex gestures and contextual interactions, especially in scenarios like gaming and video conferencing.

Remarks. Gesture recognition for earables is rapidly evolving, driven by increasing hardware sophistication and creative sensing strategies. Research trends are shifting from simple IMU-based detection to robust user-independent recognition, exploring multimodal sensor fusion, and minimizing latency and energy consumption. Additionally, personalized gesture sets and on-device learning will play a critical role in ensuring high usability across diverse users and contexts. We summarized these works in the supplementary material.

6.2 Head Motion Tracking

Head tracking enables intuitive and immersive human-computer interaction (HCI), particularly in applications such as virtual reality (VR), augmented reality (AR), and the Metaverse. Wearable devices such as VR headsets provide precise pose tracking but are typically bulky and expensive. To overcome these limitations, recent work has explored earables as a platform for head tracking, capitalizing on their natural head placement and increasing market penetration of commercial off-the-shelf (COTS) earphones. Among these, IMU-based tracking is the most straightforward and cost-effective option but suffers from cumulative drift over time. To address this issue, FaceOri [168] leverages ultrasonic ranging to estimate head pose using commercial earphones. By emitting inaudible chirps from a smartphone and measuring time-of-flight to the microphones embedded in earphones, the system could estimate head yaw and pitch. Building on this idea, HeadTrack [81] tackled practical challenges of wireless earbuds, such as limited bandwidth and asynchronous clocks, and enhanced robustness through signal processing innovations. Pushing further, IA-Track [80] proposed a hybrid sensing approach by fusing IMU-based motion tracking with acoustic-based calibration. This technique corrected accumulated IMU drift via inaudible smartphone-emitted tones, combining the responsiveness of IMUs with the stability of acoustic ranging, and showcasing improved long-term tracking reliability. All these methods operate without external infrastructure, exploiting the spatial coupling between smartphones and earphones in everyday use.

Remarks. Compared to traditional vision or headset-based methods, earphone-based tracking delivers a lightweight and scalable solution suitable for daily use. However, key challenges remain, such as handling dynamic acoustic noise, earphone misalignment, and generalization across users. Future research should explore multimodal signal fusion to boost robustness. Beyond HCI, potential applications in healthcare, cognitive state monitoring, and context-aware feedback could further elevate the value of this emerging sensing paradigm. We summarized these works in the supplementary material.

6.3 Silent Speech Recognition

Silent Speech Recognition (SSR) enables speech interaction without vocalization by detecting articulatory movements such as those of the lips, jaw, or tongue. Compared to conventional ASR, SSR provides stronger noise robustness, enhanced privacy, and accessibility for users with speech impairments. Early SSR systems leveraged ear-mounted sensors like accelerometers, proximity sensors, and cameras to detect cheek, ear canal, or tongue deformations [35, 99, 148], but they often relied on intrusive hardware or user-specific calibration, limiting their scalability. To address these limitations, recent efforts turn to commodity earables equipped with built-in microphones and IMUs, enabling SSR without dedicated hardware. These systems fall into two main categories: those based on ear canal deformation and those leveraging temporomandibular joint (TMJ) motion. We summarized these works in the supplementary material.

Ear Canal Deformation-based. This line of work detects articulation-induced changes in ear canal geometry using reflected inaudible ultrasound signals. Systems like EarCommand [91], EarSSR [159], and ReHEarSSE [48] follow a shared principle but explore different aspects of usability and generalization. EarCommand combines ultrasound sensing with motion-triggered activation via IMUs, offering energy efficiency but with a limited vocabulary and moderate enrollment needs. EarSSR focuses on adaptability, introducing incremental learning that expands the lexicon over time without forgetting prior words. ReHEarSSE emphasizes open-set recognition and phoneme-level modeling, allowing generalization to unseen words. While effective, these systems rely on continuous acoustic probing, which may interfere with audio playback or increase power consumption, posing challenges for practical deployment.

Temporomandibular Joint Movement-based. TMJ-based approaches track jaw movement patterns that directly correspond to phoneme formation, using either active acoustic sensing or inertial measurement. HPSpeech [197] adopts ultrasound to detect TMJ motion, but like ear canal methods, it raises concerns about energy use and media interference. In contrast, MuteIt [153] and Unvoiced [154] introduce passive IMU-based designs that enhance privacy and efficiency. MuteIt decouples jaw and head motion using twin IMUs, enabling phoneme-level decoding and thus scalable vocabulary construction. Unvoiced further integrates with standard ASR pipelines by converting IMU sequences into mel spectrograms for LLM-based text inference, offering seamless pipeline compatibility and eliminating the need for vocabulary-specific classifiers. Hybrid designs like QuietSync [156] and magnetic-skin systems [47] expand sensing depth by combining IMUs with ExG electrodes or magnetometers, improving robustness across users, languages, and speaking styles. These systems also reduce personalization effort, requiring only minimal calibration samples to adapt to new users.

Remarks. SSR has evolved from intrusive, rigid setups toward scalable, phoneme-aware, and integration-friendly solutions. Earlier systems relied on word-level classification, but phoneme-based decoding (e.g., MuteIt [153], Unvoiced [154]) enables open-vocabulary recognition and compatibility with standard ASR pipelines. Additionally, multi-sensor fusion and lightweight personalization (e.g., QuietSync's [156]) further enhance usability, robustness, and adaptability, marking significant strides toward practical silent speech interaction.

6.4 Live Speech Recognition and Voice Interaction

Modern earables increasingly support voice-based interaction, enabling hands-free tasks such as virtual assistant invocation, and command recognition. As these devices evolve into standalone speech interfaces, researchers have revisited core challenges around liveness detection, speaker isolation, and microphone-free operation, shifting away from conventional air-conducted input toward body-coupled and occlusion-based sensing. We summarized these works in the supplementary material.

Several systems address liveness verification to defend against spoofing or replay attacks. VibLive [108] captures both air- and bone-conducted speech to isolate body-coupled signals that are harder to forge. This approach improves robustness against ambient playback by exploiting the anatomical specificity of bone conduction. Similarly, EarSpy [29] uses the earbud's inertial sensors to detect subtle vibrations in the ear canal caused by live speech. Unlike acoustic-only methods, these systems leverage the physical connection between body and device, marking a shift toward body-coupled modalities for authentication. However, motion-based sensing also introduces potential privacy trade-offs, as movement signals could leak behavioral or contextual information.

Other systems aim to ensure that only the wearer's voice is captured. EAROE [68] introduces a body-channel voice interface that relies on occlusion-induced in-ear vibrations, filtering out external voices while preserving internal articulation. This design treats the user's body as a natural filter and improves privacy by physically attenuating environmental noise. It also employs model architectures that can reconstruct intelligible speech from narrowband internal signals, bridging the gap between physical isolation and audio fidelity. Extending this principle, EarVoice [38] explores voice activation on earphones without microphones. By exploiting acoustic patterns generated through the occlusion effect at the speaker, it detects speech and verifies speaker identity using only the playback hardware. This approach eliminates the need for a dedicated microphone and demonstrates the feasibility of ultra-minimal voice interfaces.

Remarks. Together, recent earable systems mark a transition from traditional air-conducted voice capture toward body-coupled sensing and speaker-specific isolation. Instead of indiscriminately recording ambient audio, systems like VibLive [196] and EarSpy [29] ground speech input in physiological signals, such as bone conduction or occlusion-induced vibrations, to verify liveness and suppress spoofed inputs. Approaches like EAROE [68] and EarVoice [38] physically isolate the wearer's voice, enhancing privacy and reducing false triggers even in shared or noisy environments.

6.5 Speech Enhancement

With the widespread adoption of wireless earbuds, users increasingly rely on them for telephony and conferencing in noisy, mobile environments. Unlike wired earphones with microphones near the mouth, wireless earbuds capture speech via voice pickup units (VPUs) near the ear, resulting in attenuated and noise-prone signals. At the same time, modern earables incorporate multiple microphones and auxiliary sensors, creating opportunities for advanced, context-aware speech enhancement under real-world constraints. We summarized these works in the supplementary material.

Microphone-based Approaches. Conventional beamforming provides lightweight enhancement but often fails in complex acoustic settings. Recent deep learning methods improve quality but risk exceeding the limited resources of earbuds. To bridge this gap, ClearBuds [34] employs synchronized binaural recordings for spatial beamforming, while ClearSpeech [122] and EarSpeech [67] fuse in-ear and out-ear microphones to exploit the spectral complementarity of air- and body-conducted signals. These single-earbud systems highlight practical fusion without hardware modification. Complementing them, NeuralAids [87] introduces a hardware-software co-design that integrates a speech AI accelerator with an optimized dual-path network, enabling fully on-device, low-latency speech enhancement within strict power budgets. Together, these systems illustrate the ongoing trade-off between enhancement quality and real-time, resource-efficient deployment.

Multimodal Approaches. To improve robustness under extreme noise and motion, multimodal strategies integrate bone conduction, inertial motion, or vision. In-Ear-Voice [151] embeds a bone-conduction sensor for reliable low-power voice activity detection, while VibVoice [71] fuses air and vibration signals using simulated training data to compensate for dataset scarcity. Extending this idea, Heitkaemper et al. [74] leverage bone-conducted signals to guide streaming separation for voice assistants, reducing bandwidth needs without sacrificing recognition accuracy. Beyond acoustics, Look Once to Hear [164] introduces gaze-driven enhancement, using a short visual glance for implicit speaker enrollment, aligning processing with natural attention. These methods collectively show how auxiliary modalities can strengthen speech capture when microphones alone fall short.

Beyond Enhancement. Recent work also pushes beyond traditional denoising toward augmented listening. Spatial Speech Translation [36] explores binaural hearables that not only suppress interference but also translate multiple speakers into the user's native language while preserving spatial cues and voice characteristics. By combining source separation, localization, and translation with binaural rendering, this approach extends enhancement into real-time, immersive communication support.

Remarks. Recent advances in earable speech enhancement reflect a clear shift from conventional single-microphone pipelines toward multimodal, resource-aware, and context-adaptive architectures. For example, recent systems [34, 164] emphasize lightweight, low-latency designs suitable for real-time deployment on earbuds. To address the difficulty in paired data collection, works like ClearSpeech [122] and VibVoice [71] use simulation and cross-domain transformation to ease data requirements. Additionally, some systems fuse non-acoustic signals (e.g., bone conduction, motion) to isolate the target speaker to improve robustness in noisy environments.

6.6 Facial Expression

Facial expressions serve as a rich medium for conveying emotion, intent, and user state, making them central to applications in affective computing, healthcare, and human-computer interaction. Recently, earables have emerged as a promising platform for detecting facial muscle movements through acoustic, inertial, and physiological signals captured in and around the ear. These approaches offer unobtrusive, low-power, and privacy-preserving solutions for continuous facial expression tracking. Pioneering this direction, ID.EARS [17] measures brainwave signals (EEG) and accompanying noise from the ear to recognize five facial gestures. EarIO [112] introduced active acoustic sensing via frequency-modulated continuous wave (FMCW) signals to detect skin deformations caused by facial muscle movement. Building on this, IMUFace [192] showed that subtle ear motions resulting from facial expressions could be captured using embedded IMU sensors. Their system reconstructed 3D facial landmarks in real time with only 58 mW power usage, offering a passive and energy-efficient alternative to active acoustic sensing. Futami *et al.* [54] demonstrates that infrared distance sensors embedded in various ear accessories can robustly capture skin movement around the ear to recognize facial gestures, achieving over 95% single-device accuracy. To further enhance spatial resolution and continuity, EARFace [198] utilized in-ear ultrasonic reflections combined with a transformer model, achieving continuous facial landmark estimation with sub-2 mm accuracy. This work also validated the comfort and feasibility of integrating such sensing into everyday earbuds. Exploring an alternative sensing modality, PPGface [40] demonstrated that facial muscle activity modulates blood flow, detectable via in-ear photoplethysmographic (PPG) signals. Their system achieved emotion classification accuracy exceeding 93%, even under real-world conditions, showcasing a novel physiological pathway for emotion recognition without relying on motion or acoustic data. Yuma *et al.* [195] demonstrates that PPG sensors embedded in hearable devices can also reliably recognize mouth gestures, enabling hands-free, low-cost interaction without additional hardware. By measuring the tympanic membrane temperature (TMT), Furukawa *et al.* [53] shows that right–left TMT asymmetry provides unique, brain-related physiological cues that enable above-chance emotion recognition across valence and discrete categories.

Remarks. The evolution of earable-based facial expression recognition reflects a shift from bulky, privacy-invasive systems to compact, multimodal, and real-world-ready solutions. Key challenges ahead include low-latency on-device inference, modeling of subtle expressions, and adaptability to diverse anatomical structures. We summarized these works in the supplementary material.

6.7 Wireless Earphone Location Tracking

Wireless earphones, originally designed for audio playback and communication, have evolved into promising platforms for spatial interaction. By tracking the earphones' positions and movements in real time, users can perform intuitive interactions such as mid-air gesture input and cursor control, unlocking new possibilities for hands-free interaction in AR/VR and smart home. Traditionally, earphone tracking can be achieved by having the smartphone emit inaudible acoustic signals that are received by the earphones to estimate relative distance or orientation [60]. However, accurately tracking *wireless* earphones poses significant technical challenges. First, due to BLE (Bluetooth Low Energy) protocol constraints, many systems face limitations on sampling rate and audio compression, making it hard to adopt traditional acoustic ranging techniques. Second, the clock offset between the earphone and the host device introduces non-negligible drift, which severely degrades acoustic ranging performance over time. To address these limitations, MagSound [167] introduces a dual-modality approach that combines acoustic ranging with magnetic field sensing. The key insight is that commercial earphones contain strong built-in magnets for sound production. By measuring the magnetic field strength using a smartphone's magnetometer, MagSound can predict the earphone's relative position free from clock offset. Experiments demonstrate that MagSound achieves millimeter-level 2D tracking accuracy and significantly improves handwriting recognition compared to acoustic-only baselines. BLEAR [60] further tackles a fundamental limitation of wireless earphone to enable location tracking: BLE's low sampling rate and lossy audio compression. To enable BLE earphones to receive inaudible signals ($\geq$ 17 kHz), BLEAR designs a novel frequency conversion scheme based on the nonlinearity of microphones. By mixing a high-frequency mask signal with the beacon signals, the system down-converts ultrasonic information into BLE-compatible bands.

Remarks. Recent systems explicitly embrace the limitations of commercial off-the-shelf (COTS) wireless earphones, pushing earphone tracking-based interaction toward real-world deployment. Looking forward, cross-modal learning (*e.g.*, combining magnetic, acoustic, IMU, and BLE RSS features) may further enhance robustness and generalizability. Additionally, low-power embedded implementations will be key to enabling continuous tracking on energy-constrained earable devices. We summarized these works in the supplementary material.

6.8 User Intention, Usability, and Contextual Perception

Earable devices offer a unique vantage point for capturing rich multimodal signals, especially acoustic cues from both the environment and the human body. Recent research has begun to explore how such devices can move beyond passive listening toward understanding, *i.e.*, inferring user intent, behavioral states, or environmental context from audio signals. This shift toward "earable-based user intention and contextual understanding" demands robust sensing, real-time signal interpretation, and localization under diverse and noisy real-world conditions. Takawale *et al.*[161] use in-ear acoustic sensing to detect the muscle vibrations in the ear canal and to classify periods of auditory attention. Similarly, Quan *et al.* [139] utilize the in-ear acoustic reflections to predict cognitive load levels. GrooveMeter [107] proposed a system that detects vocal and motion reactions to music via earbuds. Using in-ear microphones and IMUs, it identifies behaviors such as humming, singing along, or head-nodding, demonstrating how earables can infer user engagement from subtle auditory cues. To enable earables to attend to meaningful sounds while suppressing distractions, Veluri *et al.* [163] proposed Semantic Hearing, which extracts target sound classes (*e.g.*, sirens, alarms) from binaural input, preserving spatial cues. It supports 20 sound classes and generalizes well across users and acoustic scenes, enabling programmable auditory focus and situational awareness. EarEOG [103] uses periauricular electrodes integrated into a headphone form factor and validated that horizontal eye movements can be reliably tracked, establishing earEOG as a promising unobtrusive alternative for wearable gaze sensing. DeepEar[189] leverages binaural microphones with artificial ears to perform multi-source sound localization. Inspired by the human auditory system, it fuses handcrafted features (*e.g.*, ITD) with latent representations via a multi-sector neural network. DeepBSL[50] achieves personalized 3D sound localization by collecting a small set of user-specific HRTFs using in-ear microphones and a mobile phone.

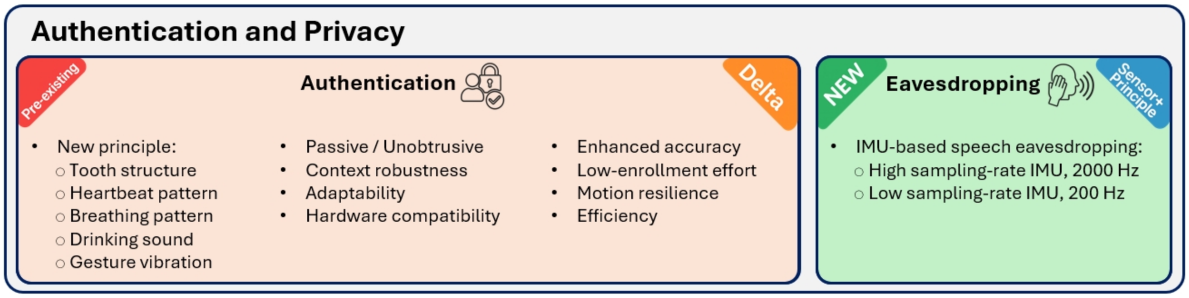

Fig. 7. Summary of recent advancements in earable authentication and privacy.

BlinkBud [113] builds a low-power earbud-camera system that performs sampled monocular 3D detection to accurately detect rear-approaching hazards despite head motion and resource constraints. To understand the health implications of everyday earphone use, Alkhalifah *et al.* [13] links prolonged wear, high listening volumes, and device sharing to increased bacterial contamination, dermatologic symptoms, and elevated risk of hearing impairment.

Remarks. These studies exemplify how earables are evolving from passive listening devices to intelligent agents capable of real-time perception and understanding. Future systems should integrate personalization pipelines to capture individual auditory characteristics (*e.g.*, HRTFs, vocal styles) for holistic context awareness. Moreover, on-device deployment of deep audio models will be essential for privacy and latency. Enabling continuous, user-centered understanding from sound will be key to the next generation of earable intelligence. We summarized these works in the supplementary material.

7 AUTHENTICATION AND PRIVACY

This section presents recent advancements in earable authentication and privacy. As illustrated in Figure 7, these developments span new biometric principles, practical improvements in system deployment, and emerging privacy threats such as IMU-based eavesdropping.

7.1 Authentication

As earables increasingly capture sensitive data, ensuring secure and seamless user authentication becomes crucial. Earlier works laid the foundation using anatomical or physiological traits, but faced challenges in usability, robustness, and real-world deployability. Recent systems improve along several key dimensions, including model efficiency, enrollment effort, passive operation, noise robustness, and adaptability, with diverse principles. We summarized these works in the supplementary material.

Ear Shape. Ear shape offers a stable anatomical feature for authentication due to its uniqueness and proximity to the device. Visual approaches like EarAuthCam [127] embed miniature cameras into earbuds to capture auricle contours, but remain sensitive to occlusion and lighting, limiting consistency in daily use. In contrast, acoustic alternatives such as Amesaka *et al.*[15] use air leakage patterns caused by audio playback, but require the user to press the device. OnePiece [57] and Gao *et al.*[55] improve usability by supporting hands-free operation and reducing interference via ultrasonic modulation. Notably, these systems can operate using standard outward-facing microphones available on commercial earables, without requiring hardware modification or custom sensor integration, which improves their deployability. However, their active probing mechanisms may still interfere with concurrent music playback, and often require calibration or adaptation to device-specific acoustic layouts and microphone placements. This reflects a broader trade-off in ear-shape-based methods: while newer systems reduce user effort and improve reliability, they frequently retain dependencies on acoustic stimulation, with implications for power efficiency, audio compatibility, and cross-device generalization.

Ear Canal Geometry. The ear canal has been extensively exploited for authentication, as its geometry uniquely shapes acoustic or sensor responses. Early acoustic probing methods [58, 124] emitted tones that disrupted playback and limited comfort. Later systems reduced this overhead: MetaEar [32] advances this direction by adopting a dedicated FMCW-based acoustic design together with a lightweight learning pipeline, enabling direct deployment on resource-constrained earable devices. Hu *et al.*[79] derived transfer functions from everyday audio captured by in-ear and out-ear microphones, while LR-Auth [78] introduced dual scenario-specific templates with a cosine similarity matcher, achieving low latency and energy-efficient operation. Yasuhara *et al.*[193] propose a bilateral ear-canal authentication approach that exploits inter-ear acoustic responses to jointly characterize left–right ear canal shape, achieving improved accuracy over single-ear methods, but relying on a dedicated acoustic pipeline that remains sensitive to device fit and environmental conditions. More recently, EarLock [16] employs sound leakage captured by external microphones to infer both ear canal and auricle shape, offering compatibility with diverse hearable types under noisy and mobile scenarios. In parallel, EarID [82] leverages ear canal scanning (ECS) but replaces black-box classifiers with a learning-free pipeline using masked cepstrum features, thereby shortening enrollment, reducing storage, and enhancing security. Beyond acoustics, EarCapAuth [69] explores capacitive sensing to reconstruct a 3D canal profile, avoiding acoustic interference but at the cost of hardware complexity. Taken together, these works reveal a trend toward passive, resource-efficient, and noise-robust authentication, while balancing accuracy, security, and hardware feasibility.

Skull and Teeth. Bone and dental structures offer deeper physiological features that remain invariant over time. Systems like MandiPass [115] and HCR-Auth [73] exploit skull vibration paths during vocalization or chirp-based probing to extract identity. While HCR-Auth expands the modality space to bone-conduction headphones, it suffers from audible probing and poor robustness under motion. In contrast, ToothSonic [173] and TeethPass [178] authenticate users via sounds generated during dental occlusion. ToothSonic supports multiple gestures for richer features, though it raises enrollment burden and gesture complexity. TeethPass reduces user effort by allowing arbitrary occlusion actions at the cost of stability. Together, these methods show that skeletal traits can enhance security and spoofing resistance, but must balance behavioral effort and system responsiveness.

Heartbeat and Breathing. Physiological signals such as heartbeats and respiration offer a passive, continuously available biometric stream. EarPass [110] adapts PPG sensing to the ear canal, improving signal stability over wrist-based setups. HeartPrint [25] bypasses optical sensing by capturing in-ear phonocardiograms using microphones, and further incorporates continual learning to adapt to temporal changes in heart rhythm without explicit re-enrollment. BreathSign [66] captures bone-conducted respiratory acoustics, enabling fast authentication with strong replay resistance. These systems are inherently noise-resilient due to the occlusion effect and well-suited for continuous verification. However, they must compensate for performance degradation under physical motion or changes in physiological state, motivating designs that combine motion-awareness with adaptive modeling.

Behavioral-based. Behavioral biometrics leverage everyday gestures that are user-controllable and socially natural. SipDeep [9] authenticates users via swallowing sounds, providing strong spoofing resistance without requiring explicit gestures. EarSlide [174] captures skin friction during finger sliding across the face, offering high discriminability through personalized acoustic features. BudsAuth [171] uses IMUs to detect micro-vibrations from habitual touch gestures and achieves very low error rates. These methods highlight a trend toward flexible, intuitive input modalities that support active verification. However, they typically require user participation and may not generalize to background or passive use unless combined with complementary sensing streams.

Multi-modal Voiceprint. Traditional voiceprint systems are vulnerable to spoofing attacks. To address this, recent works fuse vocal content with body-generated signals that reflect how speech is produced. F2Key [49] measures facial muscle dynamics via ultrasound reflections; EarPPG [41] uses PPG to detect vascular shifts during articulation; Jawthenticate [155] and PiezoBud [109] capture jaw- and skin-borne vibrations with IMUs and piezoelectric sensors. These designs emphasize spoof resistance by binding speech to physiology. They also vary in hardware cost and integration complexity—PiezoBud, for instance, achieves sub-100ms latency with minimal enrollment, but requires fine-grained vibration sensing at high sampling rates. Multi-modal designs represent a clear evolution toward resilient, low-latency, and context-adaptive speech authentication.

Remarks. Authentication methods for earables have evolved from lab-bound, intrusive setups to more practical and adaptive designs: (1) Deep models improve accuracy but often incur latency and energy costs; PiezoBud [109] mitigates this with lightweight architectures for real-time use. (2) Many systems require burdensome enrollment, such as repeated gestures [171] or stillness [25]; future designs may adopt few-shot learning or self-adaptive models to reduce this effort. (3) To improve robustness under noise and motion, some systems [25, 66] exploit occlusion effects, while others [78, 79] use ambient noise as a passive probe.

7.2 Earphone Eavesdropping

Modern earphones equipped with motion sensors can capture subtle vibrations caused by speaking, including bone conduction vibrations (BCVs) and ear canal dynamic motions (ECDMs) [28]. These signals can leak sensitive information such as spoken content, speaker identity, and gender without using a microphone [56]. Because motion sensors often don't require explicit permissions, they pose serious privacy risks as side channels. EarSpy [28] leverages high-frequency motion sensor data (up to 2 KHz) to recognize live speech from the user. They separate speech signals from noise such as body movement and earphone playback, enabling accurate, user-independent speech recognition. Gao *et al.* [56] explore a more realistic scenario with motion sensors limited to 200 Hz. They show that even with lower sampling rates, speech, identity, and gender can still be inferred using raw data and a channel attention mechanism. Their model performs well across environments and speaking volumes, highlighting the persistent risk even under system-imposed constraints.

Remarks. As motion sensors become increasingly sensitive and ubiquitous in consumer devices, the risk of side-channel eavesdropping demands urgent attention from both researchers and manufacturers. Future research should focus on developing practical defenses against motion sensor-based eavesdropping. This includes designing OS-level monitoring to detect abnormal sensor usage and exploring learning-based countermeasures that can detect or distort potential leakage. At the same time, it's important to balance privacy with legitimate sensing applications, such as health tracking or accessibility features, to ensure innovation continues safely. We summarized these works in the supplementary material.

8 RESOURCES FOR FUTURE EARABLE RESEARCH

Recent advances in earable sensing have sparked the development of a diverse ecosystem of supporting resources that enable researchers to prototype new ideas, train machine learning models, and ensure real-world robustness.

8.1 Hardware

8.1.1 Hardware Platforms. In recent years, a growing number of research-driven earable platforms have emerged, offering distinct combinations of sensing modalities, openness, and extensibility. Most were introduced after 2022, reflecting renewed interests in earable hardware (with eSense from 2018 as an early example). While prior surveys [143] reviewed application trends, they did not systematically cover platform evolution. Table 1 summarizes representative designs, and we highlight key features of selected platforms below.

eSense [96] was one of the first in-ear research platforms supporting real-time multimodal sensing. It integrates an IMU, in-ear microphone, and Bluetooth, enabling head gesture detection, proximity sensing, and audio streaming via open APIs.

Table 1. Summary of recent research-driven earable hardware platforms.

Platform (Year)	Trans-mission	Sensors	Memory	Battery Life	Special Features
eSense (2018) [96]	Bluetooth Classic	6-axis IMU (1), Microphone (1)	None	1.2 h	Real-time streaming, compact form factor
ClearBuds (2022) [34]	Custom BLE	Microphone	1 GB	40 h	Stereo recording, sub-64 μs sync error
EarAce (2023) [26]	Wi-Fi, Bluetooth	Stereo Microphones (2)	microSD	Varies	ANC control, motion artifact reduction
OpenEarable 1.3 (2022) [144]	Bluetooth BLE	6-axis IMU (1), In-ear Ultrasonic Mic (1), Pressure (1) / Temp Sensor (1)	microSD	10 h	Modular design, ultrasonic ear canal profiling
OpenEarable 2.0 (2025) [145]	Bluetooth BLE Audio	9-axis IMU (1), 3-axis Ear Canal IMU (1), Mics (3), PPG (1), Temp (1) / Pressure Sensor (1)	microSD	8 h	Multi-sensor integration, BLE audio, two ExG variants available, on-device ML
OmniBuds (2024) [132]	Bluetooth BLE Audio	9-axis IMU (1), PPG (1), Temp Sensor (1), Microphones (3)	1 GB Flash + 8 MB RAM	8 h	Cuffless blood pressure estimation, on-device ML (CNN accelerator)

ClearBuds [34] features custom synchronized earbuds that form a binaural microphone array. It supports real-time dual-channel voice capture with sub-µs sync, enabling mobile beamforming and speech enhancement, and operates for long durations using low-power transmission protocols.

EarAce [26] augments commercial ANC earphones with a plug-in module featuring stereo codecs, impedance-based fit detection, and support for ANC tuning. It enables dual-ear acoustic sensing, local storage, and configurable wireless streaming.

OpenEarable 1.3 [144] is a low-cost, open-source platform using off-the-shelf parts and a modular ear-hook design. It includes IMU, ultrasonic microphone, pressure and temperature sensors, and supports BLE-based streaming for up to 10 hours.

OpenEarable 2.0 [145] extends version 1.3 with richer sensing: in-ear, out-ear, and bone-conduction microphones, dual IMUs, pulse oximeter, and DSP-based neural inference. It includes modular expansion ports and variants supporting EEG/EMG sensing.

OmniBuds [132] integrates multi-modal sensing (IMU, PPG, temperature, microphones) with onboard inference via a CNN accelerator and co-processor. Designed as a symmetric dual-ear platform, it supports blood pressure estimation, long-term sensing, and high programmability.

All listed platforms support raw data access and offer custom APIs for algorithm prototyping and pipeline customization. Over time, we observe growing trends in multimodal integration (e.g., IMU, audio, PPG), on-device storage, and real-time inference support, signaling a maturation of earable research infrastructure for applications ranging from health monitoring to interaction sensing.

8.1.2 Earable Form Factors in Prior Studies. Prior earable research has explored a diverse set of device form factors, each offering distinct physical coupling mechanisms, sensing affordances, and practical trade-offs. The most commonly adopted form factor is in-ear earables, which are particularly well-suited for acoustic sensing using in-ear microphones. By sealing the ear canal, in-ear designs can leverage the occlusion effect to enhance low-frequency body-conducted sounds [157], making them effective for applications such as breathing monitoring, speech-related sensing, and physiological signal acquisition. In practice, ear tip materials further influence signal quality: foam tips generally provide stronger occlusion and more stable acoustic coupling than silicone tips, albeit at the cost of reduced comfort and durability [24]. As illustrated in Figure 8 (a),(b), many existing in-ear sensing systems still rely on customized earbuds or earphones, often adopting wired connections combined with external codecs to ensure high-fidelity data capture and precise synchronization—particularly for applications that require multi-channel configurations (*e.g.*, left–right or in-ear–out-ear microphone pairs).

Beyond in-ear designs, bone-conduction earables are frequently adopted when direct ear-canal occlusion is unnecessary or undesirable. As shown in Figure 8 (c), these devices couple vibrations through the skull or surrounding tissue, enabling sensing or playback without blocking the ear canal. Although bone-conduction earables do not benefit from the low-frequency amplification provided by a sealed ear canal, the propagation and reflection of low-frequency vibrations through cranial structures can still encode information related to head or skull characteristics [73]. In addition, their placement and relatively stable attachment make them well-suited for IMU-based sensing of vibration-driven gestures, jaw motion (e.g., temporomandibular joint activity), and head movements [153, 154].

Several studies further explore skin-contact earables around the ear, including electrodes placed on peri-auricular skin or in the ear canal, as illustrated in Figure 8 (d),(e). These designs are primarily motivated by electrophysiological sensing needs, such as EEG or bio-impedance measurement, where direct and stable skin contact is essential for signal fidelity. While such configurations can achieve high-quality signals, they typically require careful placement and user-specific adjustment [88, 94]. In practice, these systems often involve customized designs, and determining the optimal number and placement of electrodes remains an open research challenge [20].

Closely related to these approaches are in-ear PPG-based designs, where optical sensors are embedded within the ear canal, as shown in Figure 8 (f). These systems are commonly used for cardiovascular monitoring or for capturing subtle ear-canal deformations induced by speech or intentional interactions. Similar to electrode-based designs, the placement of PPG sensors requires careful consideration to ensure stable skin contact and effective optical isolation from ambient light, both of which are critical for reliable signal acquisition [46].

More exploratory work extends earables into ear-adjacent accessories, such as earrings or decorative wearables (Figure 8 (g)), prioritizing unobtrusiveness and social acceptability. These designs often demand careful circuit layout and mechanical integration to balance aesthetic considerations with sensing functionality, and are therefore typically limited to lightweight electronics and coarse-grained sensing tasks.

8.1.3 Hardware Realization Strategies. In addition to form-factor diversity, prior studies differ substantially in how earable hardware is realized and adapted for research purposes. At the simplest level, some works directly employ commercial off-the-shelf earables without modification, leveraging built-in sensors and default data access interfaces. While this approach offers excellent accessibility and reproducibility, it is typically constrained by limited access to raw signals, fixed sensor configurations, and closed firmware or APIs.

A second category involves modified commercial earables, where researchers adapt existing consumer devices to overcome limitations in wireless bandwidth, sampling rate, or data accessibility. Such modifications commonly include rerouting sensor outputs, bypassing on-device processing, or exporting raw signals through wired connections for external codec-based acquisition. While this approach enables high-fidelity, multi-channel data collection beyond the constraints of standard Bluetooth pipelines, it requires hardware and firmware expertise.

A third category relies on research-oriented earable platforms, as summarized in Section 8.1.1, which provide integrated sensing, synchronized data acquisition, and programmable interfaces. These platforms significantly lower the barrier to prototyping but often impose fixed sensor placements and form factors, and their multi-sensor integration may not align with all application needs.

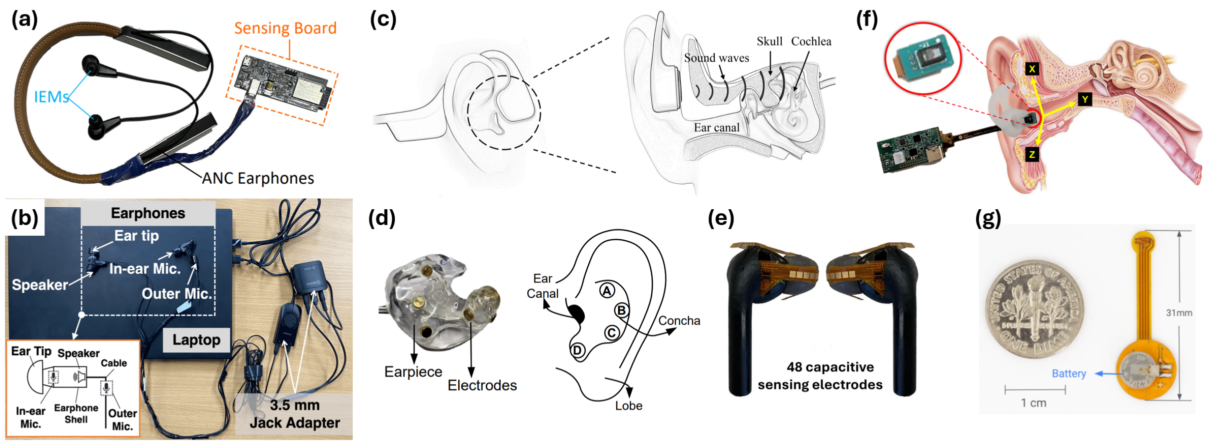

Fig. 8. Representative earable device form factors explored in prior studies, including (a,b) in-ear earables with in-ear and out-ear microphones [25, 67], (c) bone-conduction earables [73], (d,e) skin-contact earables with peri-auricular or in-ear electrodes [69, 88], (f) in-ear PPG-based designs [41], and (g) ear-adjacent accessory form factors [183]. Images are adapted from prior work.

Finally, a number of studies adopt fully self-built earable prototypes, incorporating custom mechanical designs, 3D-printed enclosures, and bespoke PCB layouts. This approach is typically motivated by the need for novel form factors, unconventional sensor placements, or application-specific constraints that cannot be met by existing devices. Although self-built prototypes offer maximal flexibility, they demand substantial engineering effort and careful validation to ensure that their sensing characteristics remain representative of realistic earable use.

Overall, research-oriented earable platforms and modified commercial earables often provide a more practical balance between engineering flexibility and deployment realism. By enabling access to raw and synchronized sensor data while retaining wearable form factors closer to everyday use, these approaches lower development overhead and improve ecological validity. Looking forward, research-facing system-level support that exposes low-level sensing interfaces on commercial earables could further reduce barriers to practical earable research.

8.2 Datasets

Datasets are essential for advancing algorithmic and learning-based earable research. Yet, unlike vision or speech data, earable signals are deeply affected by user anatomy, wearing variation, motion, and device heterogeneity, which substantially limits generalizability, reuse, and cross-study comparison. As a result, the practical utility of earable datasets depends not only on signal quality, but also on the availability of clear metadata, including data accessibility, preprocessing, and reproducibility support. To address this, several datasets have been curated in recent years, covering a range of sensing modalities and application domains. Table 2 summarizes seven representative examples, and we briefly outline their design scope and contributions below.

FatigueSet. FatigueSet [93] is a multi-modal dataset for studying mental fatigue via in-ear and head-worn sensing. It includes EEG, in-ear PPG and IMU, facial videos, and task performance logs from 29 participants alternating between resting and cognitively demanding tasks. Aligned sensor streams and self-reported fatigue scores support analysis of fatigue progression. Baseline models show that in-ear PPG and IMU signals alone can indicate fatigue, enabling passive mental state monitoring. The dataset is publicly available at https://www.kaggle.com/datasets/tanjemahamed/mental-fatigue-level-detection-fatigueset-data/data.

Table 2. Summary of representative earable datasets, including sensing modalities, number of subjects, data duration or sample counts, application scenarios, availability of raw and/or processed data, benchmark code availability, and licensing information.

Dataset	Modality	Subs.	Duration/Samples	Scenarios	Data Type	Code	Lic.
FatigueSet [93]	EEG, PPG, IMU, video	29	2 × 40 min sessions	Mental fatigue monitoring (MSIT task)	Raw	×	MIT
EarSet [131]	PPG, IMU	30	16 motion tasks per subject	PPG motion artifacts under facial, head, body movements	Raw	×	CC 4.0
EarSAVAS [200]	In-ear and out-ear Mic, IMU	42	44.5 hours	Subject-aware vocal activity (8 actions, noisy and quiet conditions)	Raw+ Processed	×	CC 4.0
OESense [123]	In-ear Mic, IMU	31	52k steps, 465 min HAR, 20,880 gestures	Step counting, HAR, face gestures	Raw	×	NA
EarGate [52]	In-ear Mic	31	52,046 steps (26,023 cycles)	Gait-based authentication under 8 walking conditions	Raw	×	NA
ClearSpeech [121]	In-ear and out-ear Mic	20	9.4 hours clean & synthetic noisy data	Speech enhancement using paired in-ear and out-ear recordings	Raw	×	NA
XRF V2 [105]	Wi-Fi CSI, IMU	40	16 hours	Cross-device activity recognition, daily activity summarization	Raw	×	NA

EarSet. EarSet [131] provides motion-robust in-ear physiological data for artifact detection and signal enhancement. Using earbuds with multi-wavelength PPG and IMU, it captures signals from 30 participants across structured facial, head, and full-body motions. The dataset includes synchronized PPG/IMU streams, task labels, and derived heart rate estimates, supporting motion-aware sensing research. The dataset is publicly available at https://zenodo.org/records/8142332.

EarSAVAS. EarSAVAS [200] is a subject-aware dataset for vocal activity sensing on earables. It contains dual-microphone and IMU data from 42 participants performing activities like speaking, coughing, and chewing, with clearly attributed speaker labels. The dataset includes both quiet and noisy conditions and supports speech-driven and privacy-aware applications. A baseline multimodal model achieves high classification accuracy. The dataset is publicly available at https://www.kaggle.com/datasets/earsavas/earsavas-dataset.

OESense. OESense [123] collects in-ear microphone and accelerometer signals for motion-resilient sensing using the occlusion effect. It includes step counting, activity recognition, and hand-to-face gesture data from 31 participants. The dataset supports tasks such as activity classification and robust gesture recognition under motion interference, and is publicly released on Kaggle at https://www.kaggle.com/datasets/dongma878/oesense.

EarGate. EarGate [52] supports gait-based authentication using inward-facing in-ear microphones. It includes walking data from 31 participants under varying ground and footwear conditions, annotated with participant ID and walking context. The dataset enables exploration of acoustic gait recognition and includes baseline models. The dataset is publicly available at https://www.kaggle.com/datasets/dongma878/eargate.

ClearSpeech. ClearSpeech [121] provides dual-microphone earbud recordings for speech enhancement. It includes clean and noisy sentence-level speech from 20 participants, with simulated noisy in-ear signals for training denoising models. The dataset supports in-ear enhancement research and is available on Kaggle at https://www.kaggle.com/datasets/dongma878/clearspeech.

XRF V2. XRF V2 [105] is a cross-device dataset combining Wi-Fi CSI with IMU signals from phones, watches, earbuds, and glasses. The earbud modality includes 9-axis IMU recordings from 40 participants performing daily activities such as walking, sitting, and gesturing. Data are time-synchronized across devices and provided in standard CSV format. The dataset supports applications in cross-device activity recognition, summarization, and multimodal fusion. The dataset is publicly available at https://www.kaggle.com/datasets/laptype/xrf-v2.

Remarks. While recent datasets reveal growing sophistication in earable sensing, broader adoption and reuse remain limited due to several persistent challenges. Unlike image or speech domains, earable data is shaped by limitations from anatomical sensing, diverse modalities, and custom hardware. Specifically, (1) Contextual fragility: Earable signals vary with user anatomy, wearing conditions, motion, and environment, reducing consistency and generalizability. (2) Cross-dataset compatibility: Differences in sensor placement, sampling rates, and labeling hinder multimodal integration and often require task-specific data collection. (3) Reproducibility barriers: Custom hardware and undocumented mounting variations limit reimplementation, even with similar components. (4) Annotation granularity: Subtle behaviors like chewing or soft speech lack clear standards, making label quality and evaluation inconsistent. To support scalable and robust research, future datasets should emphasize standardization, open hardware, multimodal synchronization, and rich annotations.

8.3 Earable Signal Quality Assessment Tools

While datasets provide the foundation for algorithm development, the real-world reliability of earable sensing ultimately hinges on signal quality. In practice, signals captured by earables, such as in-ear audio, PPG, EEG, or IMU data, are highly susceptible to degradation from poor ear canal sealing, motion artifacts, and inconsistent sensor contact. However, many systems implicitly assume clean input, overlooking the dynamic and user-dependent variations in signal quality that occur during daily use. These degradations can significantly impair sensing accuracy and lead to unreliable or misleading downstream inferences in applications such as health monitoring, neural decoding, or attention tracking. As such, *assessing and ensuring signal quality is not merely an optimization step, but a critical prerequisite for building robust and generalizable earable systems.*

Demirel *et al.* [46] addressed a frequently overlooked factor: the quality of ear canal sealing, which affects both in-ear microphone and photoplethysmography (PPG) signals. They proposed a novel, hardware-free method for estimating air leakage using the distortion patterns captured by in-ear microphones. By modeling acoustic distortion and applying machine learning, their system could detect poor fitting conditions in real time, offering a practical solution for unobtrusively monitoring ear fit and consequently improving sensing reliability. Jayas *et al.* [88] tackled the challenge of ensuring signal fidelity in ear-EEG systems, where body movements often disrupt electrode-skin contact. They proposed an in-situ channel selection method based on manifold learning, which extracts geometric features (dimension and curvature) from EEG signal manifolds to distinguish valid EEG channels from noise-contaminated ones. Kim *et al.* [101] proposes a real-time earphone wearing-detection method that replaces conventional proximity sensors with PPG-based sensing and lightweight edge-AI k-NN classification, achieving F1 scores above 0.95. Complementing these efforts, Berent *et al.* [20] conducted a large-scale comparison of ear EEG with traditional scalp and intracranial EEG using clinical data. They demonstrated that ear EEG recorded during sleep exhibited higher signal quality than during wakefulness, with coherence and spectral similarities to scalp and intracranial signals, especially in low-frequency bands (delta, theta, alpha). Notably, contralateral ear EEG channels provided better signal resemblance than unilateral ones, confirming their relevance for high-quality neural monitoring. Pazuelo *et al.* [136] also conducted a study that provides a comprehensive validation of a dry-electrode in-ear EEG device, showing moderate-to-high correlations with temporal scalp EEG across alpha rhythms, eye-movement artifacts, and NREM sleep features, demonstrating its feasibility for comfortable, unobtrusive brain monitoring outside clinical settings.

Remarks. Signal quality assessment for earables is important for real-world deployment of ear-centric sensing systems. Future work should aim to develop unified signal quality metrics covering EEG, PPG, acoustic, and IMU signals simultaneously. Moreover, integrating signal quality assessment into closed-loop systems can enable dynamic reconfiguration of sensing parameters (*e.g.*, gain, filtering, sensor fusion) in response to degradation. The inclusion of personalized signal quality baselines and cross-device generalizability remains an open challenge. Additionally, quality-aware machine learning models that explicitly model uncertainty due to signal degradation may improve robustness in health and cognitive inference tasks. We summarized these works in the supplementary material.

9 RESEARCH GAPS AND CHALLENGES

Despite the rapid progress in earable-based sensing, existing research remains fragmented across sensing modalities, system designs, and application scenarios. By synthesizing insights from the surveyed literature, we identify several critical research gaps and challenges that currently limit the robustness, scalability, and real-world applicability of earable-based sensing systems.

Signal Modality and Sensing Quality. A fundamental gap lies in the mismatch between sensing modalities and the specific requirements of earable-based sensing. While a variety of open-source and proprietary earable platforms have been proposed [132, 145], most systems rely on custom-built hardware with off-the-shelf sensors that are not explicitly optimized for earable form factors or physiological sensing tasks. In practice, many sensors offer limited configurability (e.g., fixed gain settings [42]), lack robust wearing or fitting detection [46], and are sensitive to placement variations [154]. As a result, although signals are technically available, they are often weak, noisy, or unstable, which significantly complicates downstream signal processing and modeling. This sensing-level fragility propagates through the entire system pipeline, placing an undue burden on learning-based models to compensate for hardware limitations.

Generalization and Robust Modeling. Although prior work has demonstrated promising performance across a diverse set of physiological indicators, most existing approaches operate under constrained assumptions. Common restrictions include requiring users to remain stationary, perform a single activity, avoid speech, or undergo explicit per-user or per-device calibration. These assumptions limit generalizability across users, devices, and contexts, posing a major barrier to real-world deployment. Improving robustness requires addressing fundamental modeling challenges: how to effectively represent heterogeneous, noisy physiological signals; how to fuse complementary modalities and devices in a principled manner; and how to achieve population-level generalization without repeated calibration. Current approaches often treat these challenges independently, rather than adopting holistic system-level modeling strategies.

Application Scope and "Earable-Unique" Value. Many existing earable-based applications largely replicate functionalities already provided by traditional wearables, such as heart rate estimation [24], activity recognition [152], or calorie tracking. While earables can serve as viable alternatives, such substitution alone may not justify widespread adoption. To unlock the full potential of earables, future systems must target applications that are difficult or impossible to achieve with other form factors. As the wearable devices closest to the head, earables are uniquely positioned to sense signals from the upper respiratory system [117], cranial vasculature, brain activity [18], and the oral cavity [195]. Emerging examples include dietary monitoring, toothbrushing analysis [173, 185, 186], spirometry-like respiratory assessment [30, 177, 179], and auditory assistance. However, a central challenge remains: how to systematically understand and model the relationship between subtle signal variations and the underlying specific physiological processes they reflect.

Data Availability, Benchmarking, and Reproducibility. Progress in this area is further hindered by the lack of standardized datasets and benchmarking protocols. Most studies rely on data collected using custom hardware configurations and task-specific ground truth [76, 109], making cross-study comparison and reproducibility difficult. Moreover, existing datasets are typically small-scale and short-term, often collected through brief field studies or controlled experiments [131, 200]. The absence of longitudinal, in-the-wild datasets limits our understanding of long-term variability, user adaptation, and real-world robustness. Without shared benchmarks and consistent evaluation methodologies, it remains challenging to assess true progress across the field.

Real-Time Performance and System Integration. Despite the growing number of demonstrated applications, real-time performance remains a significant challenge. Many systems stream raw or minimally processed data wirelessly to smartphones for inference, while computationally intensive models may further offload processing to the cloud. This architecture introduces latency, increases energy consumption, and raises usability concerns, particularly for high-bandwidth modalities such as audio. Wireless bandwidth constraints, compression-induced signal degradation, and limited on-device computational resources all complicate the goal of truly real-time, untethered operation. Achieving practical deployment requires rethinking data transmission strategies and computational partitioning across earables, smartphones, and edge devices [132].

Energy Efficiency and Power Constraints. Energy remains a primary bottleneck for earable-based sensing because the tiny form factor of these devices severely limits battery capacity. Many applications require continuous sensing, and the combined cost of on-device processing/inference and wireless transmission can quickly drain the battery, degrading usability and user experience. Key challenges in reducing energy consumption include optimizing end-to-end signal processing and modeling pipelines, dynamically adjusting sensing and transmission parameters (e.g., sampling rate and duty cycling), and designing adaptive offloading strategies that balance accuracy, latency, and energy cost. Exploring complementary approaches such as opportunistic sensing [199] and head-motion-based energy harvesting [64] also remains largely underexplored. A key challenge lies in practically integrating and exploiting energy harvesting in earables – not only to prolong battery lifetime, but also to realise energy-neutral operation and move towards battery-less earable platforms

Compatibility with Commercial Earables. Most existing research assumes dedicated research prototypes, with limited consideration of the realities of commercial earable usage. In practice, earables are primarily used for music playback, voice calls, and noise control (e.g., active noise cancellation and transparency modes), all of which can interfere with sensing signals [42]. Additionally, commercial earables exhibit substantial diversity in form factors, including in-ear, over-ear, open-ear, and ear-clip designs. Techniques that rely heavily on in-ear acoustics may not directly transfer across these variants. Designing sensing and inference methods that coexist with everyday usage patterns and heterogeneous hardware remains an open challenge.

Ethics, Privacy, and Long-Term Use. Finally, ethical and privacy considerations pose persistent challenges for earable-based sensing. Long-term wearability, potential discomfort, and unintended interference with auditory perception must be carefully evaluated. Moreover, earables often capture sensitive audio and health-related signals, raising concerns about data security, informed consent, and user trust. Compared to technical challenges, these issues have received relatively limited attention in the literature, yet they are essential for responsible and sustainable deployment of earable-based systems.

These gaps and challenges together highlight that advancing earable-based sensing requires coordinated progress across hardware design, sensing algorithms, system integration, and user-centered considerations. Understanding these limitations provides critical context for future research directions. In the following section, we explore potential solutions to address these challenges spanning hardware, software, application design, energy optimization, and usability.

10 FUTURE RESEARCH DIRECTIONS

We conducted a chronological analysis of prior earable publications and point out several open research gaps and challenges. In response to the question raised in the introduction: *earable computing is far from being a saturated research area. On the contrary, we believe it is entering an exciting new phase*. In the following sections, we outline five key research directions, **Hardware, Application, Software, Energy Efficiency, and Usability**, that call for further investigation in this rapidly evolving domain.

10.1 Hardware

While Section 8.1.1 has introduced a range of open-source earable platforms, practical deployment in real-world settings still poses significant hardware-related challenges. This section focuses on two core challenges: how to efficiently transmit high-fidelity sensor data, and how to support intelligent, low-latency processing through on-device AI inference.

Wireless Data Transmission. Modern earables predominantly rely on Bluetooth connections, offering convenience by eliminating cables and enabling seamless smartphone integration. However, traditional Bluetooth protocols introduce several limitations for sensing applications. First, Bluetooth Classic typically supports only single-channel audio transmission at 16 kHz, which restricts applications such as ultrasound-based sensing [49, 197] that require higher sampling rates (e.g., 48 kHz). Although the newer LE Audio [3] standard improves capabilities, supporting dual-channel microphone streaming at 16 kHz per channel and potentially 48 kHz if fully implemented, it still depends on manufacturer support and remains insufficient for demanding applications like dual-channel ultrasound. Second, applications relying on multi-channel audio [34, 67, 122] or multimodal fusion [41, 49] often face bandwidth bottlenecks, particularly when continuous raw data streaming is needed. To address these challenges, one potential solution is to design hybrid transmission mechanisms, allowing the earable to dynamically switch between Bluetooth and wired modes depending on the application's bandwidth demands. Specifically, high-bandwidth applications could leverage wired connections to ensure data fidelity and real-time throughput, while Bluetooth can be reserved for low-data-rate scenarios, prioritizing user mobility and comfort. Alternatively, future designs may offload buffered sensor data to the charging case during recharging, where higher-bandwidth links (like Wi-Fi) might be available for bundled transmission to the cloud.

Research-oriented Sensing Interfaces. Progress in mobile and wearable sensing depends not only on novel algorithms, but also on the availability of standardized, low-level interfaces on commercial devices that support development, data collection, and large-scale experimentation. However, as discussed in Section 8.1.3, the wearable domain currently lacks a unified, plug-and-play sensing interface that can expose stable low-level signals on commercial off-the-shelf devices. Most commercial wearables rely on proprietary audio processing pipelines like ANC, provide restricted APIs, and operate under bandwidth constraints, all of which limit access to raw sensing data. As a result, prior studies often rely on heterogeneous device setups, ranging from modified commercial earables to fully self-built prototypes, along with diverse data collection protocols. In practice, several partial solutions have emerged. For example, EarAce [26] leverages the ANC hardware in commercial earables to expose intermediate acoustic signals, enabling more flexible and versatile acoustic sensing without fully rebuilding the device pipeline. Open-source wearable platforms [132, 145] offer an alternative by exposing synchronized multimodal signals and programmable sensing interfaces, thereby lowering the barrier to prototyping. Nevertheless, these platforms are often less optimized than commercial wearables in terms of comfort, fit, and long-term wearability, which limits their suitability for large-scale or in-the-wild deployments. Looking forward, research-facing system-level support that selectively exposes low-level sensing interfaces on commercial wearables could substantially improve accessibility.

AI Inference Capability. Recent platforms like OpenEarable 2.0 [145] have begun to integrate on-device AI accelerators to enable local inference, reducing reliance on external computation and improving real-time responsiveness. Specifically, OpenEarable incorporates a Tensilica HiFi 3z DSP Core for running neural network models at the edge, reflecting a broader trend toward more intelligent and autonomous processing in earables. However, current accelerators are mainly optimized for lightweight models with limited memory and computational budgets, and more demanding tasks, such as high-fidelity audio enhancement or large-scale speech recognition, still require model simplification or off-device computation. Looking forward, advances in low-power

AI chip technologies may allow parts of the sensing pipeline, such as feature extraction or early-stage inference, to be performed directly on earables. This shift can reduce wireless bandwidth usage, lower latency, and mitigate privacy risks by limiting raw data transmission. To be effective in earable settings, however, such chips must support ultra-low idle power, efficient streaming and multi-modal inference, and tight integration with on-device memory and sensing pipelines. Consequently, on-device intelligence is likely to complement rather than replace off-device computation, leading to hybrid architectures that balance efficiency, capability, and practicality.

10.2 Application

As sensing hardware and algorithms improve, earables are enabling new application domains that extend beyond the capabilities of traditional wearables. This section focuses on two key directions: applications uniquely enabled by the ear's anatomy, and the shift from basic physiological metrics to richer, clinically meaningful biomarkers for long-term health monitoring.

Earable-exclusive Applications. While many earable applications have replicated functions from other wearables, such as heart rate monitoring or activity recognition, tasks uniquely suited to the ear remain underexplored. The ear's anatomy, however, presents two promising sensing directions. First, its proximity to the brain and major blood vessels makes it an ideal site for neurophysiological sensing. Ear-EEG enables applications like sleep monitoring, mental load estimation, emotion recognition, and seizure detection. Yet, this comes with trade-offs: more electrodes improve spatial resolution but reduce comfort and increase design complexity, risking poor fit or signal instability [94]. Second, being near the respiratory tract and vocal apparatus, earables can capture breathing-related signals. They support applications like respiratory rate tracking, breathing mode classification, and lung function assessment, all without intrusive equipment like chest straps or nasal cannulas. In short, earables unlock access to physiological signals that are difficult to measure elsewhere, enabling a new class of exclusive applications. Despite these advancements, broader and novel applications remain worthy of exploration.

Advanced Biomarker Monitoring. Most existing earable applications focus on basic physiological metrics, such as heart rate for cardiovascular monitoring or breathing rate and mode for respiratory tracking. However, advancing toward deeper, clinically relevant biomarkers could transform earables from lifestyle accessories into credible health-monitoring tools. In cardiovascular assessment, for example, [33] has demonstrated that PPG can be used not only to track pulse rate but also to extract waveform-derived features such as rise time, pulse transit time, and pulse wave velocity. These indicators reflect arterial compliance and stiffness, key parameters in assessing vascular aging and identifying early signs of cardiovascular disease. In respiratory health, continuous monitoring of breath sounds using in-ear microphones may enable the detection of conditions such as asthma, chronic obstructive pulmonary disease, pneumonia, and sleep apnea. Acoustic features like wheezes, crackles, and apneic pauses serve as non-invasive markers of respiratory dysfunction. By facilitating passive, long-term data collection in daily life, earables could support early detection of respiratory disorders, reducing the need for frequent clinical visits and enabling more proactive health management.

10.3 Software

Beyond hardware, software plays a critical role in enabling robust and intelligent earable sensing. Modern systems increasingly depend on algorithms to fuse multimodal signals, coordinate across devices, and extract meaningful inferences. This section highlights key software directions, multi-sensor fusion, cross-wearable collaboration, and advanced modeling techniques that support scalable and adaptive earable applications.

Multi-sensor Fusion. Integrating multiple sensing modalities within earable systems offers software-level opportunities to improve sensing accuracy, motion resilience, and physiological coverage. Through data fusion algorithms, complementary signals can be combined to perform cross-validation and artifact mitigation. For example, IMU data can help verify motion patterns and be used to suppress motion-induced noise in PPG or in-ear microphone signals via signal filtering or regression techniques. Beyond denoising, fusion pipelines can unlock new capabilities: recent work [162] demonstrated that jointly analyzing PPG and in-ear microphone signals enables non-invasive blood pressure estimation by measuring vascular transit time, which leverages the different propagation speeds of acoustic and blood signals. Despite these benefits, fusion poses software challenges, including cross-modal synchronization, adaptive fusion strategies, and efficient processing frameworks. These must be carefully addressed to operate within the resource constraints of earable systems.

Cross-wearable Collaboration. Earables provide privileged access to both internal physiological signals and environmental context, but other wearable devices, such as smartwatches, smart glasses, and head-mounted displays, offer complementary sensing perspectives. While each platform has specific strengths, they also suffer from inherent limitations when used in isolation: smartwatches primarily capture arm and wrist motion but struggle with core posture or internal signals; smart glasses can detect eye and facial movements but are less effective for deep physiological monitoring; and head-mounted displays can track head pose or gaze but are bulky and energy-intensive. Cross-wearable collaboration offers a path to overcome these individual limitations by leveraging synchronized sensing across multiple devices. For instance, combining earables with smartwatches can enable more precise full-body activity or pose estimation; pairing with smart glasses allows for joint audio-gaze interaction and spatial awareness; integrating with AR headsets supports spatial audio and immersive context-aware communication. Enabling such cooperative sensing across wearables will require robust time synchronization, efficient data fusion protocols, and energy-aware communication strategies designed for heterogeneous, resource-constrained platforms.

Advanced Data Modeling Techniques. Precise sensing in earable applications relies not only on hardware but also heavily on effective modeling techniques. Early works primarily used classical signal processing methods and traditional machine learning models tailored to specific tasks. More recently, self-supervised learning and transfer learning have enabled better feature extraction from unlabeled data and cross-modal adaptation, improving robustness under real-world conditions. A growing trend is the adoption of foundation models for physiological sensing. Instead of training separate models for each application, researchers increasingly use large pretrained models that can be adapted or fine-tuned for new tasks with limited data. For example, PaPaGei [138] introduces a foundation model for PPG, pretrained on over 57,000 hours of data, and shows strong performance across various health tasks. These developments suggest that modular, reusable models and pretraining strategies could significantly accelerate progress in earable sensing. Future work should explore multi-modal foundation models, domain-adaptive pretraining (across audio, IMU, PPG), and lightweight model architectures optimized for on-device inference.

10.4 Energy Efficiency

Energy efficiency remains a fundamental constraint for earable devices due to their limited size and battery capacity. This section outlines how sensor selection, computation offloading strategies, and power asymmetry across earbuds influence energy consumption, and discusses possible design strategies to extend battery life while preserving functionality.

Sensor and Algorithm Selection. Many earable applications can be achieved with different sensing principles and modalities. For example, in facial gesture recognition, IMU-based approaches [202] generally consume less power but may offer lower accuracy compared to microphone-based [191] solutions. However, each sensor type comes with its sampling rate, algorithmic complexity, and hardware demands. Some applications require sensor arrays or multimodal fusion, all of which can significantly increase energy consumption. In academic research, even for the same application, optimization goals often vary: some prioritize accuracy, others focus on robustness to noise, and still others aim to minimize power or latency. In real-world scenarios, however, developers must carefully navigate these trade-offs. When a slight reduction in accuracy is acceptable, it may be more practical to choose a lower-power modality or a simplified algorithm to extend battery life and improve long-term usability.

Therefore, a thorough investigation and benchmarking across different sensing modalities are necessary before a given application can be efficiently deployed in real-world settings.

Computation Offloading Strategy. Due to the inherent limitations of earables, including restricted memory, limited processing power, and small battery capacity, many practical systems offload computation to smartphones or cloud servers. While this enables more complex models and reduces on-device load, it introduces several challenges. First, continuous wireless transmission (especially of audio or high-frequency signals) can significantly increase power consumption. Second, offloading introduces latency, which may hinder real-time applications. Third, and most importantly, transmitting raw physiological data, such as breathing sounds, EEG, or in-ear audio, raises serious privacy concerns, as such data can reveal sensitive health or behavioral information. Although systems often adopt empirical rules for deciding when to offload, no standard framework exists. Recent efforts have explored lightweight on-device pre-processing, such as feature extraction or anonymization, to reduce both data volume and privacy risk. Building adaptive strategies that balance performance, efficiency, and privacy based on context remains an open challenge for future work.

Battery Balance. As previously discussed, most Bluetooth earphones use an asymmetric architecture where one side, typically the primary earbud, handles microphone input and communicates with the smartphone, while the secondary connects via a low-power internal link such as BLE [12]. This simplifies communication but often causes faster battery drain on the primary side. A similar imbalance occurs in on-device computation, where tasks may be assigned to one earbud or unevenly distributed based on sensor location or system design. While sufficient for simple applications, this becomes problematic for binaural tasks like beamforming, spatial localization, or EEG asymmetry analysis, where early depletion of one side can impair function. To address this, some systems [84] explore dynamic role-switching and workload balancing based on battery status or sensing needs, or support task migration to the better-powered side. Still, maintaining battery symmetry remains an open challenge, especially in long-duration, high-demand scenarios.

Energy Harvesting and Energy-Neutrality. Recent advances in miniaturized energy harvesting [61], including harvesting from body motion, heat, and even sound, create new opportunities to embed harvesters into earables and reduce reliance on finite battery capacity. Beyond power supply, an emerging direction is to exploit harvested-energy signals as a sensing modality [98, 106]: energy harvesters worn on the body can produce information-rich electrical signatures tied to user motion and context, potentially replacing (or duty-cycling) conventional sensors and thereby cutting power draw, footprint, and system complexity of the wearable device. Together, these directions open a pathway toward energy-neutral earable systems and, ultimately, battery-less earable operation.

10.5 Usability

Beyond technical performance, the success of earables depends heavily on how well they integrate into users' daily lives. This section evaluates usability considerations including long-term hearing aid usage, the influence of audio playback features such as ANC and transparency mode, and the design trade-offs of open-ear versus in-ear form factors, all of which impact comfort, compliance, and sensing robustness.

Hearing Aid Usability. Earables encompass a range of devices, including earbuds, hearing aids, and bone-conduction headphones. Yet most research has focused on earbuds, with hearing aids remaining relatively underexplored in the sensing community. Despite this, hearing aids offer unique advantages that make them highly promising for long-term physiological and behavioral monitoring. Unlike earbuds, hearing aids are typically used throughout the day [5], making them well-suited for continuous sensing. This persistent wear supports fine-grained tracking of vital signals, activity patterns, and long-term health trends. Many users are elderly or hearing-impaired, and could particularly benefit from context-aware applications such as fall detection, cognitive monitoring, hearing environment adaptation, or social engagement assessment. Hearing aids may also enable real-time speech clarification, hearing training, and voice-based reminders to assist with daily living.

However, hearing aids are often highly personalized, customized to each user's audiometric profile, ear canal shape, and comfort preferences. This process, involving hearing threshold fitting and acoustic sealing [11], can hinder the deployment of general-purpose sensing systems. Earable applications for hearing aids should therefore be designed to reduce enrollment burden and support generalizability across diverse user populations.

Impact of ANC and Transparency Mode. Active Noise Cancellation (ANC) and Transparency Mode have become standard features in modern earables, allowing users to adapt their auditory experience across different environments. However, these audio processing modes can significantly interfere with sensing applications that rely on playback signals and recorded echoes, such as those used for speech enhancement, authentication, or respiratory monitoring. For instance, ANC may introduce frequency-dependent colorations or phase shifts, while Transparency Mode may alter echo profiles due to signal mixing from internal and external microphones. These modifications can distort the original audio cues or feedback patterns critical for sensing accuracy. Despite their prevalence, it does not remain easy to characterize or compensate for such effects in practice. This is due in part to the proprietary nature of ANC/transparency implementations, which are not publicly documented, and also to the use of varying algorithms across device models, often dictated by microphone placement and vendor-specific designs. As a result, applications involving audio feedback or echo analysis must explicitly account for these transformations, though doing so remains an open and underexplored challenge.

Open-ear Form Factor. Most early earable devices adopt an in-ear, occlusive form factor. However, prolonged ear canal occlusion may pose health risks [126], including increased humidity, microbial growth, and earwax buildup. In response, open-ear designs, such as bone-conduction [4] and ear-clip [1] earphones have emerged, offering improved breathability and long-term comfort. Yet, this shift introduces new challenges. Many physiological and activity-sensing applications leverage the occlusion effect to enhance low-frequency body-conducted sounds. Open-ear designs, by leaving the ear canal open, lose this acoustic amplification advantage. Additionally, the altered positioning and looser fit of open-ear devices complicate sensor deployment: for instance, PPG sensors may lose skin contact, and IMUs on ear-clip headphones may be more susceptible to motion artifacts due to increased movement. Due to these limitations, only a small number of existing studies [73] have specifically targeted sensing with open-ear earphones. This highlights the significant design and engineering barriers that must be addressed before such devices can support a wider range of physiological or behavioral sensing applications.

11 CONCLUSION

This survey presents a comprehensive update on the state of earable computing from 2022 to 2025. By analyzing over one hundred recent publications, we systematically compare new findings with prior efforts, highlighting key advancements in sensing principles, system performance, hardware integration, and deployment feasibility. Despite these promising developments, earable computing still faces critical challenges around hardware usability, energy efficiency, and real-world generalizability, which demand thoughtful system design and user-centric engineering to ensure seamless daily use. Looking ahead, we believe earables are poised to become a cornerstone of next-generation ubiquitous computing. With their unique ability to passively and continuously capture both internal physiological signals and external contextual cues, earables represent a rich sensing platform for intelligent, human-centered applications. By addressing the remaining gaps and exploring underutilized opportunities, future research can further elevate earables from peripheral accessories to essential components of health monitoring, interaction, and ambient intelligence systems.

Acknowledgments

This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 2 grant (Fund Code: T2EP20124-0046), as well as the EPSRC grant HearFit (EP/S023046/1).

References

[1] Online. Bose Ultra Open Earbuds. https://www.bose.com/p/earbuds/bose-ultra-open-earbuds/ULT-HEADPHONEOPN.html.

[2] Online. Heart Rate Monitoring - Powerbeats Pro 2 - Beats by Dre. https://www.beatsbydre.com/sg/heart-rate-monitoring.

[3] Online. LE Audio: The next generation of Bluetooth® audio. https://www.bluetooth.com/learn-about-bluetooth/feature-enhancements/le-audio/.

[4] Online. OpenRun Pro2 – Shokz. https://shokz.com/pages/openrunpro2.

[5] Online. Should you wear hearing aids all the time? https://www.healthyhearing.com/report/53141-Wear-hearing-aids-all-the-time.

[6] Online. True Wireless Stereo Earbuds Global Market Report 2025. https://www.thebusinessresearchcompany.com/report/true-wireless-stereo-earbuds-global-market-report.

[7] Online. Why are teenagers dependent on headphones? https://oshkoshnorthstar.org/4858/columns/why-are-teenagers-dependent-on-headphones.

[8] Yilizati Abulaiti, Shoko Kimura, and Guillaume Lopez. 2024. Earable-Based Posture Inference and Correction System for Deskwork Behavior Transformation. In *Congress of the International Ergonomics Association*. Springer, 503–509.

[9] Awais Ahmed, Panlong Yang, Adeel Feroz Mirza, Taha Khan, Muhammad Rizwan, Ammar Hawbani, Miao Pan, and Zhu Han. 2024. SipDeep: Swallowing-Based Transparent Authentication via Bone-Conducted In-Ear Acoustics. *IEEE Transactions on Mobile Computing* (2024).

[10] Tousif Ahmed, Md Mahbubur Rahman, Ebrahim Nemati, Mohsin Yusuf Ahmed, Jilong Kuang, and Alex Jun Gao. 2023. Remote breathing rate tracking in stationary position using the motion and acoustic sensors of earables. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–22.

[11] Nasim Alamdari, Edward Lobarinas, and Nasser Kehtarnavaz. 2020. Personalization of hearing aid compression by human-in-the-loop deep reinforcement learning. *IEEE access* 8 (2020), 203503–203515.

[12] Alexander Goldin. 2025. Smart Dynamic Role Switching (SDRS) in Bluetooth TWS Earbuds: Seamless Connectivity & Better Voice Clarity. https://www.linkedin.com/pulse/smart-dynamic-role-switching-sdrs-bluetooth-tws-earbuds-goldin-ycwxf/. Accessed: 2025-04-23.

[13] Azzam Alkhalifah, Amal Hussein Mohammed, Norah Ahmad Alswoailem, Mohsina Huq, Sultan Alenazi, Mishaal Abdallah Alrshed, Hatem Mohammed Alturki, Ahmad Abdulaziz Almatroudi, Hajed Obaid Alharbi, and Basmah F Alharbi. 2025. Earphone use habits and their association with auditory and dermatologic complications. *Scientific Reports* 15, 1 (2025), 41848.

[14] Khaled Alkiek, Moustafa Youssef, and Khaled A Harras. 2023. EarBender: Enabling rich imu-based natural hand-to-ear interaction in commodity earables. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing*. 333–338.

[15] 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 ACM International Symposium on Wearable Computers*. 119–123.

[16] Takashi Amesaka, Hiroki Watanabe, Yuta Sugiura, Masanori Sugimoto, and Buntarou Shizuki. 2025. EarLock: Personal Authentication System for Hearables Using Sound Leakage Signals. *IEEE Transactions on Mobile Computing* (2025).

[17] Hyunjin An, Eunkyu Oh, Yoosung Kim, Seonho Kim, Dasom Park, and Changhoon Oh. 2025. ID. EARS: One-Ear EEG Device with Biosignal Noise for Real-Time Gesture Recognition and Various Interactions. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–18.

[18] Abdul Aziz, Nhat Pham, Neel Vora, Cody Reynolds, Jaime Lehnen, Pooja Venkatesh, Zhuoran Yao, Jay Harvey, Tam Vu, Kan Ding, et al. 2025. An unobtrusive and lightweight ear-worn system for continuous epileptic seizure detection. *ACM Transactions on Computing for Healthcare* 6, 1 (2025), 1–31.

[19] Ananta Narayanan Balaji, Andrea Ferlini, Fahim Kawsar, and Alessandro Montanari. 2023. Stereo-bp: Non-invasive blood pressure sensing with earables. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*. 96–102.

[20] Jonathan Berent, Emiel Vereycken, Nigel Colenbier, Pieter van Mierlo, and Caroline Neuray. 2024. Quality Assessment And Neurophysiological Signal Resemblance Of Ear EEG. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 691–696.

[21] Sizhen Bian, Rakita Strahinja, Philipp Schilk, Clénin Marc-André, Silvano Cortesi, Kanika Dheman, Elio Reinschmidt, and Michele Magno. 2024. Earable and Wrist-worn Setup for Accurate Step Counting Utilizing Body-Area Electrostatic Sensing. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 904–910.

[22] Jacob Brown, Yang Liu, and Cecilia Mascolo. 2023. Yawning Detection using Earphone Inertial Measurement Units. In *Proceedings of the 2nd Workshop on Smart Wearable Systems and Applications*. 7–13.

[23] Kayla-Jade Butkow, Ting Dang, Andrea Ferlini, Dong Ma, Yang Liu, and Cecilia Mascolo. 2024. An evaluation of heart rate monitoring with in-ear microphones under motion. *Pervasive and Mobile Computing* 100 (2024), 101913.

[24] Kayla-Jade Butkow, Ting Dang, Andrea Ferlini, Dong Ma, and Cecilia Mascolo. 2023. heart: Motion-resilient heart rate monitoring with in-ear microphones. In *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 200–209.

[25] Yetong Cao, Chao Cai, Fan Li, Zhe Chen, and Jun Luo. 2023. HeartPrint: Passive Heart Sounds Authentication Exploiting In-Ear Microphones. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 1–10.

[26] Yetong Cao, Chao Cai, 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* 7, 2 (2023), 1–23.

[27] Yetong Cao, Huijie Chen, Fan Li, and Yu Wang. 2021. Crisp-BP: Continuous wrist PPG-based blood pressure measurement. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 378–391.

[28] Yetong Cao, Fan Li, Huijie Chen, Xiaochen Liu, Chunhui Duan, and Yu Wang. 2023. I can hear you without a microphone: Live speech eavesdropping from earphone motion sensors. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 1–10.

[29] Yetong Cao, Fan Li, Huijie Chen, Xiaochen Liu, Shengchun Zhai, Song Yang, and Yu Wang. 2023. Live speech recognition via earphone motion sensors. *IEEE Transactions on Mobile Computing* 23, 6 (2023), 7284–7300.

[30] Yetong Cao, Dong Ma, Wentao Xie, Qian Zhang, and Jun Luo. 2025. ESPIRO: Natural Pulmonary Function Monitoring via Earphone-Acquired Speech. In *Proceedings of the 31st Annual International Conference on Mobile Computing and Networking (Hong Kong, China)(MobiCom'25). Association for Computing Machinery, New York, NY, USA*. 1–16.

[31] Justin Chan, Antonio Glenn, Malek Itani, Lisa R Mancl, Emily Gallagher, Randall Bly, Shwetak Patel, and Shyamnath Gollakota. 2023. Wireless earbuds for low-cost hearing screening. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*. 84–95.

[32] Zhuo Chang, Lin Wang, Binbin Li, and Wenyuan Liu. 2022. MetaEar: Imperceptible acoustic side channel continuous authentication based on ERTF. *Electronics* 11, 20 (2022), 3401.

[33] Peter H Charlton, Birutė Paliakaitė, Kristjan Pilt, Martin Bachler, Serena Zanelli, Daniel Kulin, John Allen, Magid Hallab, Elisabetta Bianchini, Christopher C Mayer, et al. 2022. Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from VascAgeNet. *American Journal of Physiology-Heart and Circulatory Physiology* 322, 4 (2022), H493–H522.

[34] Ishan Chatterjee, Maruchi Kim, Vivek Jayaram, Shyamnath Gollakota, Ira Kemelmacher, Shwetak Patel, and Steven M Seitz. 2022. ClearBuds: wireless binaural earbuds for learning-based speech enhancement. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 384–396.

[35] Tuochao Chen, Benjamin Steeper, Kinan Alsheikh, Songyun Tao, François Guimbretière, and Cheng Zhang. 2020. C-face: Continuously reconstructing facial expressions by deep learning contours of the face with ear-mounted miniature cameras. In *Proceedings of the 33rd annual ACM symposium on user interface software and technology*. 112–125.

[36] Tuochao Chen, Qirui Wang, Runlin He, and Shyamnath Gollakota. 2025. Spatial Speech Translation: Translating Across Space With Binaural Hearables. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–19.

[37] Tao Chen, Yongjie Yang, Xiaoran Fan, Xiuzhen Guo, Jie Xiong, and Longfei Shangguan. 2024. Exploring the feasibility of remote cardiac auscultation using earphones. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 357–372.

[38] Tao Chen, Yongjie Yang, Chonghao Qiu, Xiaoran Fan, Xiuzhen Guo, and Longfei Shangguan. 2024. Enabling Hands-Free Voice Assistant Activation on Earphones. In *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*. 155–168.

[39] Hyeob Choi, Sunghwan Ji, Jinkyuk Kim, Chang Keun Jung, Jaehyung Shin, Yun Jeong Baek, Heyjin Kim, Eun Young Kim, Dae Young Kim, and Jae-Young Lim. 2025. Development and validation of an earphone-based system to measure short physical performance battery in older adults. *Experimental Gerontology* 201 (2025), 112703.

[40] 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 (2022), 1–32.

[41] Seokmin Choi, Junghwan Yim, Yincheng Jin, Yang Gao, Jiyang Li, and Zhanpeng Jin. 2023. EarPPG: Securing Your Identity with Your Ears. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 835–849.

[42] 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*. 147–153.

[43] Kenneth Christofferson, Sejal Bhalla, Joseph Cafazzo, and Alex Mariakakis. 2024. On the Production and Measurement of Cardiac Sounds in the Ear Canal. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 685–690.

[44] Kenneth Christofferson, Xuyang Chen, Zeyu Wang, Alex Mariakakis, and Yuntao Wang. 2022. Sleep sound classification using anc-enabled earbuds. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, 397–402.

[45] Kenneth Christofferson, Michelle Lin, Joseph A Cafazzo, and Alex Mariakakis. 2025. ArtEARial: Arterial Pressure Waveform Reconstruction Using Earbud Audio. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–24.

[46] Berken Utku Demirel, Ting Dang, Khaldoon Al-Naimi, Fahim Kawsar, and Alessandro Montanari. 2024. Unobtrusive air leakage estimation for earables with in-ear microphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–29.

[47] Penghao Dong, Yizong Li, Si Chen, Justin T Grafstein, Irfaan Khan, and Shanshan Yao. 2023. Decoding silent speech commands from articulatory movements through soft magnetic skin and machine learning. *Materials Horizons* 10, 12 (2023), 5607–5620.

[48] Xuefu Dong, Yifei Chen, Yuuki Nishiyama, Kaoru Sezaki, Yuntao Wang, Ken Christofferson, and Alex Mariakakis. 2024. ReHEarSSE: Recognizing Hidden-in-the-Ear Silently Spelled Expressions. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.

[49] Di Duan, Zehua Sun, Tao Ni, Shuaicheng Li, Xiaohua Jia, Weitao Xu, and Tianxing Li. 2024. F2key: Dynamically converting your face into a private key based on cots headphones for reliable voice interaction. In *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*. 127–140.

[50] Awny M El-Mohandes, Navid H Zandi, and Rong Zheng. 2023. DeepBSL: 3-D personalized deep binaural sound localization on earable devices. *IEEE Internet of Things Journal* 10, 21 (2023), 19004–19013.

[51] Xiaoran Fan, David Pearl, Richard Howard, Longfei Shangguan, and Trausti Thormundsson. 2023. Apg: Audioplethysmography for cardiac monitoring in hearables. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. 1–15.

[52] 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*. 337–349.

[53] Kaita Furukawa, Xinyu Shui, Ming Li, and Dan Zhang. 2025. Utilizing Tympanic Membrane Temperature for Earphone-Based Emotion Recognition. *Sensors* 25, 14 (2025), 4411.

[54] Kyosuke Futami, Kohei Oyama, and Kazuya Murao. 2022. Augmenting ear accessories for facial gesture input using infrared distance sensor array. *Electronics* 11, 9 (2022), 1480.

[55] Ming Gao, Jiatong Chen, Xin Tong, Ruitong Ye, Yike Chen, Fu Xiao, and Jinsong Han. 2025. Quick-pass Continuous Authentication with Real-time Biometrics Extraction on COTS Earphones Using Out-ear Microphones. *IEEE Transactions on Mobile Computing* (2025).

[56] Mengzhen Gao, Helei Cui, Yanze Xie, Yaxing Chen, Zhiwen Yu, Bin Guo, and Xingliang Yuan. 2023. Practical earphone eavesdropping with built-in motion sensors. In *2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2219–2226.

[57] Ming Gao, Xin Tong, Jiatong Chen, Yike Chen, Fu Xiao, and Jinsong Han. 2024. Eternity in a Second: Quick-pass Continuous Authentication Using Out-ear Microphones. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*. 675–688.

[58] 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 (2019), 1–24.

[59] Sandip Chakraborty Suchetana Chakraborty Garvit Chugh, Indrajeet Ghosh. 2025. BiteSense: Earable-Based Inertial Sensing for Eating Behaviour Assessment. In *IEEE PerCom*.

[60] Linfei Ge, Wentao Xie, Jin Zhang, and Qian Zhang. 2024. BLEAR: Practical Wireless Earphone Tracking under BLE protocol. In *2024 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 164–173.

[61] Taher Ghomian and Shahab Mehraeen. 2019. Survey of energy scavenging for wearable and implantable devices. *Energy* 178 (2019), 33–49. doi:10.1016/j.energy.2019.04.088

[62] Google. 2025. Google Pixel Buds. https://support.google.com/googlepixelbuds/answer/7560934?hl=en. Accessed: 2025-04-19.

[63] Google. 2025. Pixel Buds Pro 2. https://store.google.com/product/pixel_buds_pro_2?hl=en-GB. Accessed: 2025-04-19.

[64] Raffaele Guida and Tommaso Melodia. 2025. Powering Smart Wireless Implantable Medical Devices: Toward an internet of self-powered intra-body things. *Ad Hoc Networks* 169 (2025), 103748.

[65] Feiyu Han, Panlong Yang, Yuanhao Feng, Weiwei Jiang, Youwei Zhang, and Xiang-Yang Li. 2024. Earsleep: In-ear acoustic-based physical and physiological activity recognition for sleep stage detection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (2024), 1–31.

[66] Feiyu Han, Panlong Yang, Shaojie Yan, Haohua Du, and Yuanhao Feng. 2023. BreathSign: Transparent and continuous in-ear authentication using bone-conducted breathing biometrics. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 1–10.

[67] Feiyu Han, Panlong Yang, You Zuo, Fei Shang, Fenglei Xu, and Xiang-Yang Li. 2024. Earspeech: Exploring in-ear occlusion effect on earphones for data-efficient airborne speech enhancement. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–30.

[68] Feiyu Han, You Zuo, Weiwei Jiang, Dawei Yan, Yuxin Zhao, Panlong Yang, and Yubo Yan. 2025. EAROE: Enabling Body-Channel Voice Interaction Interface on Earphones via Occlusion Effect. *IEEE Internet of Things Journal* (2025).

[69] Richard Hanser, Tobias Röddiger, Till Riedel, and Michael Beigl. 2024. EarCapAuth: Biometric Method for Earables Using Capacitive Sensing Eartips. (2024).

[70] Peter Hausamann, Martin Daumer, Paul R MacNeilage, and Stefan Glasauer. 2019. Ecological momentary assessment of head motion: Toward normative data of head stabilization. *Frontiers in human neuroscience* 13 (2019), 179.

[71] Lixing He, Haozheng Hou, Shuyao Shi, Xian Shuai, and Zhenyu Yan. 2023. Towards bone-conducted vibration speech enhancement on head-mounted wearables. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*. 14–27.

[72] Yu He, Yincheng Jin, and Zhanpeng Jin. 2025. Handleap: towards contact-free gesture interaction with earphones via acoustic sensing. *CCF Transactions on Pervasive Computing and Interaction* (2025), 1–17.

[73] Zhixiang He, Jing Chen, Cong Wu, Kun He, Ruiying Du, Ju Jia, Yangyang Gu, and Xiping Sun. 2024. HCR-Auth: Reliable Bone Conduction Earphone Authentication with Head Contact Response. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–27.

[74] Jens Heitkaemper, Joe Caroselli, Max McKinnon, Arun Narayanan, and Nathan Howard. 2025. Bone Conducted Signal Guided Speech Enhancement For Voice Assistant on Earbuds. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[75] Nagisa Hidaka, Satoshi Kurose, Nana Takao, Takumi Miyauchi, Sachiko Nakajima, Sawako Yoshiuchi, Aya Fujii, Kazuhisa Takahashi, Hiromi Tsutsumi, Daiki Habu, et al. 2022. Masticatory behaviors and gender differences in people with obesity as measured via an earphone-style light-sensor-based mastication meter. *Nutrients* 14, 14 (2022), 2990.

[76] Changshuo Hu, Thivya Kandappu, Yang Liu, Cecilia Mascolo, and Dong Ma. 2024. BreathPro: Monitoring breathing mode during running with earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (2024), 1–25.

[77] Changshuo Hu, Thivya Kandappu, Jake Stuchbury-Wass, Yang Liu, Anthony Tang, Cecilia Mascolo, and Dong Ma. 2024. Detecting Foot Strikes during Running with Earbuds. In *Proceedings of the Workshop on Body-Centric Computing Systems*. 35–40.

[78] Changshuo Hu, Xiao Ma, Xinger Huang, Yiran Shen, and Dong Ma. 2024. LR-Auth: Towards Practical Implementation of Implicit User Authentication on Earbuds. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–27.

[79] Changshuo Hu, Xiao Ma, Dong Ma, and Ting Dang. 2023. Lightweight and non-invasive user authentication on earables. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*. 36–41.

[80] Jingyang Hu, Hongbo Jiang, Daibo Liu, Zhu Xiao, Qibo Zhang, Jiangchuan Liu, and Schahram Dustdar. 2023. Combining IMU with acoustics for head motion tracking leveraging wireless earphone. *IEEE Transactions on Mobile Computing* 23, 6 (2023), 6835–6847.

[81] Jingyang Hu, Hongbo Jiang, Zhu Xiao, Siyu Chen, Schahram Dustdar, and Jiangchuan Liu. 2023. HeadTrack: Real-Time Human–Computer Interaction via Wireless Earphones. *IEEE Journal on Selected Areas in Communications* 42, 4 (2023), 990–1002.

[82] Chenpei Huang, Hui Zhong, Zhen Chen, Xiaoqi Qin, Lan Zhang, Xiaoyong Yuan, and Miao Pan. 2025. Leveraging Binary Ear Canal Scanning Code for User Authentication in Earbuds. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[83] Yongzhi Huang, Jiayi Zhao, and Kaishun Wu. 2025. Reconstructing Ear Canal Channels for Fine-Grained Detection of Tympanic Membrane Changes. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–48.

[84] Jonathan D Hurwitz. 2020. Cell Age Balancing For Wireless Earbuds. (2020).

[85] Koki Iguma, Kazuya Murao, and Hiroki Watanabe. 2023. Input Interface with Touch and Non-touch Interactions using Atmospheric Pressure for Hearable Devices. In *Proceedings of the 2023 ACM International Symposium on Wearable Computers*. 1–5.

[86] Md Saiful Islam, Md Mahbubur Rahman, Mehrab Bin Morshed, David J Lin, Yunzhi Li, Hao Zhou, Wendy Berry Mendes, and Jilong Kuang. 2025. BallistoBud: Heart Rate Variability Monitoring using Earbud Accelerometry for Stress Assessment. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–21.

[87] Malek Itani, Tuochao Chen, Arun Raghavan, Gavriel Kohlberg, and Shyamnath Gollakota. 2025. Wireless hearables with programmable speech ai accelerators. In *Proceedings of the 31st Annual International Conference on Mobile Computing and Networking*. 863–877.

[88] Tanuja Jayas, R Jhanavi, Bhartendu Kumar, Vivek BS, Kartik Muralidharan, Arpan Pal, and Jayavardhana Gubbi. 2024. I Know it's EEG: An In-situ Selection of Ear-EEG channels. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 667–672.

[89] Nan Jiang, Terence Sim, and Jun Han. 2022. EarWalk: towards walking posture identification using earables. In *Proceedings of the 23rd Annual International Workshop on Mobile Computing Systems and Applications*. 35–40.

[90] Yincheng Jin, Yang Gao, Xiaotao Guo, Jun Wen, Zhengxiong Li, and Zhanpeng Jin. 2022. Earhealth: an earphone-based acoustic otoscope for detection of multiple ear diseases in daily life. In *Proceedings of the 20th annual international conference on mobile systems, applications and services*. 397–408.

[91] Yincheng Jin, Yang Gao, Xuhai Xu, Seokmin Choi, Jiyang Li, Feng Liu, Zhengxiong Li, and Zhanpeng Jin. 2022. EarCommand: " Hearing" Your Silent Speech Commands In Ear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–28.

[92] Chang Keun Jung, Jinkyuk Kim, and Hye Chang Rhim. 2023. Validation of an ear-worn wearable gait analysis device. *Sensors* 23, 3 (2023), 1244.

[93] Manasa Kalanadhabhatta, Chulhong Min, Alessandro Montanari, and Fahim Kawsar. 2021. FatigueSet: A multi-modal dataset for modeling mental fatigue and fatigability. In *International Conference on Pervasive Computing Technologies for Healthcare*. Springer, 204–217.

[94] Netiwit Kaongoen, Jaehoon Choi, Jin Woo Choi, Haram Kwon, Chaeeun Hwang, Guebin Hwang, Byung Hyung Kim, and Sungho Jo. 2023. The future of wearable EEG: A review of ear-EEG technology and its applications. *Journal of neural engineering* 20, 5 (2023), 051002.

[95] Ryan Kaveh, Carolyn Schwendeman, Leslie Pu, Ana C Arias, and Rikky Muller. 2024. Wireless ear EEG to monitor drowsiness. *Nature Communications* 15, 1 (2024), 6520.

[96] Fahim Kawsar, Chulhong Min, Akhil Mathur, and Alessandro Montanari. 2018. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing* 17, 3 (2018), 83–89.

[97] Tamisa Ketmalasiri, Yu Yvonne Wu, Kayla-Jade Butkow, Cecilia Mascolo, and Yang Liu. 2024. IMChew: Chewing Analysis using Earphone Inertial Measurement Units. In *Proceedings of the Workshop on Body-Centric Computing Systems*. 29–34.

[98] Sara Khalifa, Guohao Lan, Mahbub Hassan, Aruna Seneviratne, and Sajal K. Das. 2018. HARKE: Human Activity Recognition from Kinetic Energy Harvesting Data in Wearable Devices. *IEEE Transactions on Mobile Computing* 17, 6 (2018), 1353–1368. doi:10.1109/TMC.2017.2761744

[99] Prerna Khanna, Tanmay Srivastava, Shijia Pan, Shubham Jain, and Phuc Nguyen. 2021. JawSense: recognizing unvoiced sound using a low-cost ear-worn system. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*. 44–49.

[100] Jiwan Kim, Mingyu Han, and Ian Oakley. 2025. BudsID: Mobile-Ready and Expressive Finger Identification Input for Earbuds. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.

[101] Raehyeong Kim, Joungmin Park, Jaeseong Kim, Jongwon Oh, and Seung Eun Lee. 2025. Real-Time True Wireless Stereo Wearing Detection Using a PPG Sensor with Edge AI. *Electronics* 14, 19 (2025), 3911.

[102] Tae-Ho Kim, Dominic Jaworski, Rakesh Sethi, Elise Huisman, Kam Fung, and Edward J Park. 2025. Neckband-type earphone for continuous monitoring of cardiovascular symptoms via self-powered box knot pulse-wave sensor. *npj Biomedical Innovations* 2, 1 (2025), 4.

[103] Tobias King, Michael Knierim, Philipp Lepold, Christopher Clarke, Hans Gellersen, Michael Beigl, and Tobias Röddiger. 2025. earEOG via periauricular electrodes to facilitate eye tracking in a natural headphone form factor. *Scientific Reports* 15, 1 (2025), 32437.

[104] Michael Thomas Knierim, Lukas Braun, and Monica Perusquia-Hernandez. 2024. Warmth on Demand: Designing Headphones for Enhanced Thermal Comfort in Work Environments. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[105] Bo Lan, Pei Li, Jiaxi Yin, Yunpeng Song, Ge Wang, Han Ding, Jinsong Han, and Fei Wang. 2025. Xrf v2: A dataset for action summarization with wi-fi signals, and imus in phones, watches, earbuds, and glasses. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–41.

[106] Guohao Lan, Dong Ma, Weitao Xu, Mahbub Hassan, and Wen Hu. 2020. Capacitor-based Activity Sensing for Kinetic-powered Wearable IoTs. *ACM Trans. Internet Things* 1, 1, Article 2 (March 2020), 26 pages. doi:10.1145/3362124

[107] Euihyeok Lee, Chulhong Min, Jaeseung Lee, Jin Yu, and Seungwoo Kang. 2023. GrooveMeter: Enabling Music Engagement-aware Apps by Detecting Reactions to Daily Music Listening via Earable Sensing. In *Proceedings of the 31st ACM International Conference on Multimedia*. 7728–7736.

[108] Feng Li, Jiayi Zhao, Huan Yang, Dongxiao Yu, Yuanfeng Zhou, and Yiran Shen. 2024. VibHead: An authentication scheme for smart headsets through vibration. *ACM Transactions on Sensor Networks* 20, 4 (2024), 1–21.

[109] Gen Li, Huaili Zeng, Hanqing Guo, Yidong Ren, Aiden Dixon, Zhichao Cao, and Tianxing Li. 2024. Piezobud: A piezo-aided secure earbud with practical speaker authentication. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*. 564–577.

[110] Jiao Li, Yang Liu, Zhenjiang Li, and Jin Zhang. 2023. EarPass: Continuous User Authentication with In-ear PPG. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing*. 327–332.

[111] Jiao Li, Yang Liu, Tao Sun, Ziheng Zhou, and Jin Zhang. 2025. Earable-based Continuous Blood Pressure Monitoring via a Single-Point Flexible Sensor. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 828–833.

[112] Ke Li, Ruidong Zhang, Bo Liang, François Guimbretière, and Cheng Zhang. 2022. Eario: A low-power acoustic sensing earable for continuously tracking detailed facial movements. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–24.

[113] Yunzhe Li, Jiajun Yan, Yuzhou Wei, Kechen Liu, Yize Zhao, Chong Zhang, Hongzi Zhu, Li Lu, Shan Chang, and Minyi Guo. 2025. BlinkBud: Detecting Hazards from Behind via Sampled Monocular 3D Detection on a Single Earbud. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–26.

[114] Zhaohui Li, Wei Luo, Yongmin Zhang, Jianxi Chen, Yuanchao Shu, and Yaoxue Zhang. 2024. ASLiquid: Non-Intrusive Liquid Counterfeit Identification with Your Earphones. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*. 41–53.

[115] Jianwei Liu, Wenfan Song, Leming Shen, Jinsong Han, Xian Xu, and Kui Ren. 2021. Mandipass: Secure and usable user authentication via earphone imu. In *2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 674–684.

[116] Peichen Liu, Sadasivan Puthusserypady, I Scott MacKenzie, Cihan Uyanik, and John Paulin Hansen. 2025. EarEOG: Using Headphones and Around-the-Ear EOG Signals for Real-Time Wheelchair Control. *Proceedings of the ACM on Human-Computer Interaction* 9, 3 (2025), 1–16.

[117] Yang Liu, Kayla-Jade Butkow, Jake Stuchbury-Wass, Adam Pullin, Dong Ma, and Cecilia Mascolo. 2025. Respear: Earable-based robust respiratory rate monitoring. In *2025 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 67–77.

[118] Yang Liu and Alessandro Montanari. 2025. EarFusion: Quality-Aware Fusion of In-Ear Audio and Photoplethysmography for Heart Rate Monitoring. In *Proceedings of the 3rd ACM International Workshop on Intelligent Acoustic Systems and Applications*. 31–36.

[119] Yang Liu, Qiang Yang, Kayla-Jade Butkow, Jake Stuchbury-Wass, Dong Ma, and Cecilia Mascolo. 2025. EarMeter: Continuous Respiration Volume Monitoring with Earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–29.

[120] Shengzhe Lyu, Yongliang Chen, Di Duan, Renqi Jia, and Weitao Xu. 2024. Earda: Towards accurate and data-efficient earable activity sensing. In *2024 IEEE Coupling of Sensing & Computing in AIoT Systems (CSCAIoT)*. IEEE, 1–7.

[121] Dong Ma. 2024. ClearSpeech Dataset: Improving Voice Quality of Earbuds Using Both In-Ear and Out-Ear Microphones. https://www.kaggle.com/datasets/dongma878/clearspeech. Accessed: April 2025.

[122] Dong Ma, Ting Dang, Ming Ding, and Rajesh Balan. 2024. ClearSpeech: Improving Voice Quality of Earbuds Using Both In-Ear and Out-Ear Microphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–25.

[123] Dong Ma, Andrea Ferlini, and Cecilia Mascolo. 2021. Oesense: employing occlusion effect for in-ear human sensing. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. 175–187.

[124] Shivangi Mahto, Takayuki Arakawa, and Takafumi Koshinak. 2018. Ear acoustic biometrics using inaudible signals and its application to continuous user authentication. In *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE, 1407–1411.

[125] Carlo Massaroni, Andrea Nicolò, Daniela Lo Presti, Massimo Sacchetti, Sergio Silvestri, and Emiliano Schena. 2019. Contact-based methods for measuring respiratory rate. *Sensors* 19, 4 (2019), 908.

[126] R Mazlan, L Saim, A Thomas, R Said, and B Liyab. 2002. Ear infection and hearing loss amongst headphone users. *The Malaysian journal of medical sciences: MJMS* 9, 2 (2002), 17.

[127] Yurina Mizuho, Yohei Kawasaki, Takashi Amesaka, and Yuta Sugiura. 2024. Earauthcam: Personal identification and authentication method using ear images acquired with a camera-equipped hearable device. In *Proceedings of the Augmented Humans International Conference 2024*. 119–130.

[128] Mobvoi. 2025. Mobvoi Earbuds Gesture. https://www.mobvoi.com/ph/pages/mobvoiearbudsgesture. Accessed: 2025-04-19.

[129] Abdilahi Mohamoud, Joseph Jensen, and Kevin G Buda. 2024. Consumer-grade wearable cardiac monitors: What they do well, and what needs work. *Cleveland Clinic Journal of Medicine* 91, 1 (2024), 23–29.

[130] Vimal Mollyn, Riku Arakawa, Mayank Goel, Chris Harrison, and Karan Ahuja. 2023. Imuposer: Full-body pose estimation using imus in phones, watches, and earbuds. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–12.

[131] Alessandro Montanari, Andrea Ferlini, Ananta Narayanan Balaji, Cecilia Mascolo, and Fahim Kawsar. 2023. Earset: A multi-modal dataset for studying the impact of head and facial movements on in-ear ppg signals. *Scientific data* 10, 1 (2023), 850.

[132] Alessandro Montanari, Ashok Thangarajan, Khaldoon Al-Naimi, Andrea Ferlini, Yang Liu, Ananta Narayanan Balaji, and Fahim Kawsar. 2024. OmniBuds: A Sensory Earable Platform for Advanced Bio-Sensing and On-Device Machine Learning. *arXiv preprint arXiv:2410.04775* (2024).

[133] Dennis Moschina, Tobias Röddiger, and Michael Beigl. 2023. Vertical Jump Test Using an Earable Accelerometer. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing*. 324–326.

[134] Yugo Nakamura, Yuta Ohira, and Yutaka Arakawa. 2025. Eat2pic-Mobile: A Chew-Draw Interactive System for Encouraging Mindful Chewing with Smartphones and Earable Devices. In *International Conference on Human-Computer Interaction*. Springer, 345–364.

[135] Payod Panda, Molly Jane Nicholas, David Nguyen, Eyal Ofek, Michel Pahud, Sean Rintel, Mar Gonzalez-Franco, Ken Hinckley, and Jaron Lanier. 2023. Beyond audio: Towards a design space of headphones as a site for interaction and sensing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 904–916.

[136] Jeremy Pazuelo, Jose Yesith Juez, Hanane Moumane, Jan Pyrzowski, Liliana Mayor, Fredy Enrique Segura-Quijano, Mario Valderrama, and Michel Le Van Quyen. 2024. Evaluating the electroencephalographic signal quality of an in-ear wearable device. *Sensors* 24, 12 (2024), 3973.

[137] Bradley J Petek, Mostafa A Al-Alusi, Nathaniel Moulson, Aubrey J Grant, Cyril Besson, J Sawalla Guseh, Meagan M Wasfy, Vincent Gremeaux, Timothy W Churchill, and Aaron L Baggish. 2023. Consumer wearable health and fitness technology in cardiovascular medicine: JACC state-of-the-art review. *Journal of the American College of Cardiology* 82, 3 (2023), 245–264.

[138] Arvind Pillai, Dimitris Spathis, Fahim Kawsar, and Mohammad Malekzadeh. [n. d.]. PaPaGei: Open Foundation Models for Optical Physiological Signals. In *The Thirteenth International Conference on Learning Representations*.

[139] Jiatao Quan, Khaldoon Al-Naimi, Xijia Wei, Yang Liu, Fahim Kawsar, Alessandro Montanari, and Ting Dang. 2025. Cognitive Load Monitoring via Earable Acoustic Sensing. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[140] Md Mahbubur Rahman, Mehrab Bin Morshed, Nafiul Rashid, Sharath Chandrashekhara, and Jilong Kuang. 2025. MindfulBuddy: Extracting Comprehensive Breathing Biomarkers for Breathing Exercise Biofeedback Using Earbud Motion Sensors. *IEEE Internet of Things Journal* (2025).

[141] Pamela L Ramage-Morin, Rex Banks, Dany Pineault, and Maha Atrach. 2019. Tinnitus in Canada. *Health Rep* 30, 3 (2019), 3–11.

[142] Junbin Ren, Ruihao Zheng, Wenbo Zhang, Dong She, Yuting Bai, ZhanPeng Jin, and Yang Gao. 2025. Motion2Press: Cross Model Learning from IMU to Plantar Pressure for Gait Analysis. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–33.

[143] Tobias Röddiger, Christopher Clarke, Paula Breitling, Tim Schneegans, Haibin Zhao, Hans Gellersen, and Michael Beigl. 2022. Sensing with earables: A systematic literature review and taxonomy of phenomena. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 6, 3 (2022), 1–57.

[144] Tobias Röddiger, Tobias King, Dylan Ray Roodt, Christopher Clarke, and Michael Beigl. 2022. Openearable: Open hardware earable sensing platform. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*. 246–251.

[145] Tobias Röddiger, Michael Küttner, Philipp Lepold, Tobias King, Dennis Moschina, Oliver Bagge, Joseph A Paradiso, Christopher Clarke, and Michael Beigl. 2025. OpenEarable 2.0: Open-Source Earphone Platform for Physiological Ear Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 1 (2025), 1–33.

[146] Julia Romero, Andrea Ferlini, Dimitris Spathis, Ting Dang, Katayoun Farrahi, Fahim Kawsar, and Alessandro Montanari. 2024. OptiBreathe: An earable-based PPG system for continuous respiration rate, breathing phase, and tidal volume monitoring. In *Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications*. 99–106.

[147] RTINGS. 2025. The 5 Best Noise Cancelling Earbuds of 2025. https://www.rtings.com/headphones/reviews/best/noise-cancelling-earbuds. Accessed: 2025-04-19.

[148] Himanshu Sahni, Abdelkareem Bedri, Gabriel Reyes, Pavleen Thukral, Zehua Guo, Thad Starner, and Maysam Ghovanloo. 2014. The tongue and ear interface: a wearable system for silent speech recognition. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*. 47–54.

[149] Furrukh Sana, Eric M Isselbacher, Jagmeet P Singh, E Kevin Heist, Bhupesh Pathik, and Antonis A Armoundas. 2020. Wearable devices for ambulatory cardiac monitoring: JACC state-of-the-art review. *Journal of the American College of Cardiology* 75, 13 (2020), 1582–1592.

[150] Yukina Sato, Takashi Amesaka, Takumi Yamamoto, Hiroki Watanabe, and Yuta Sugiura. 2024. Exploring User-Defined Gestures as Input for Hearables and Recognizing Ear-Level Gestures with IMUs. *Proceedings of the ACM on Human-Computer Interaction* 8, MHCI (2024), 1–23.

[151] Philipp Schilk, Niccolò Polvani, Andrea Ronco, Milos Cernak, and Michele Magno. 2023. In-ear-voice: Towards milli-watt audio enhancement with bone-conduction microphones for in-ear sensing platforms. In *Proceedings of the 8th ACM/IEEE Conference on Internet of Things Design and Implementation*. 1–12.

[152] Thomas Sepanosian and Ozlem Durmaz Incel. 2024. Boxing Gesture Recognition in Real-Time using Earable IMUs. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 673–678.

[153] Tanmay Srivastava, Prerna Khanna, Shijia Pan, Phuc Nguyen, and Shubham Jain. 2022. Muteit: Jaw motion based unvoiced command recognition using earable. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–26.

[154] Tanmay Srivastava, Prerna Khanna, Shijia Pan, Phuc Nguyen, and Shubham Jain. 2024. Unvoiced: Designing an LLM-assisted Unvoiced User Interface using Earables. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*. 784–798.

[155] Tanmay Srivastava, Shijia Pan, Phuc Nguyen, and Shubham Jain. 2023. Jawthenticate: Microphone-free Speech-based Authentication using Jaw Motion and Facial Vibrations. (2023).

[156] Tanmay Srivastava, R Michael Winters, Thomas Gable, Yu Te Wang, Teresa LaScala, and Ivan J Tashev. 2024. Whispering wearables: Multimodal approach to silent speech recognition with head-worn devices. In *Proceedings of the 26th International Conference on Multimodal Interaction*. 214–223.

[157] Stefan Stenfelt. 2011. Acoustic and physiologic aspects of bone conduction hearing. (2011).

[158] Jake Stuchbury-Wass, Yang Liu, Kayla-Jade Butkow, Joshua Carter, Mathias Ciliberto, Quing Yang, Ezio Preatoni, Dong Ma, and Cecilia Mascolo. 2025. WalkEar: holistic gait monitoring using earables. In *2025 IEEE International Conference on Pervasive Computing and Communications (PerCom)*.

[159] Xue Sun, Jie Xiong, Chao Feng, Haoyu Li, Yuli Wu, Dingyi Fang, and Xiaojiang Chen. 2024. Earssr: Silent speech recognition via earphones. *IEEE Transactions on Mobile Computing* 23, 8 (2024), 8493–8507.

[160] Shunta Suzuki, Takashi Amesaka, Hiroki Watanabe, Buntarou Shizuki, and Yuta Sugiura. 2024. EarHover: Mid-Air Gesture Recognition for Hearables Using Sound Leakage Signals. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[161] Harshvardhan Takawale, Yang Liu, Khaldoon Al-Naimi, Fahim Kawsar, and Alessandro Montanari. 2025. Towards Detecting Auditory Attention from in-Ear Muscle Contractions using Commodity Earbuds. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[162] Hoang Truong, Alessandro Montanari, and Fahim Kawsar. 2022. Non-invasive blood pressure monitoring with multi-modal in-ear sensing. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6–10.

[163] Bandhav Veluri, Malek Itani, Justin Chan, Takuya Yoshioka, and Shyamnath Gollakota. 2023. Semantic hearing: Programming acoustic scenes with binaural hearables. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[164] Bandhav Veluri, Malek Itani, Tuochao Chen, Takuya Yoshioka, and Shyamnath Gollakota. 2024. Look once to hear: Target speech hearing with noisy examples. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.

[165] Diana Vitazkova, Erik Foltan, Helena Kosnacova, Michal Micjan, Martin Donoval, Anton Kuzma, Martin Kopani, and Erik Vavrinsky. 2024. Advances in respiratory monitoring: a comprehensive review of wearable and remote technologies. *Biosensors* 14, 2 (2024), 90.

[166] Haolong Wang, Zihao Yang, Hao Wang, and Lin Wang. 2025. Ear2Pos: A Dual-IMU Framework for Full-Body Pose Estimation Using Earbuds. *IEEE Internet of Things Journal* (2025).

[167] Lihao Wang, Wei Wang, Haipeng Dai, and Shizhe Liu. 2023. Magsound: Magnetic field assisted wireless earphone tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–32.

[168] Yuntao Wang, Jiexin Ding, Ishan Chatterjee, Farshid Salemi Parizi, Yuzhou Zhuang, Yukang Yan, Shwetak Patel, and Yuanchun Shi. 2022. Faceori: Tracking head position and orientation using ultrasonic ranging on earphones. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–12.

[169] Yang Wang, Feng Hong, Yufei Jiang, Chenyu Bao, Chao Liu, and Zhongwen Guo. 2024. Toothfairy: Real-time tooth-by-tooth brushing monitor using earphone reversed signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–19.

[170] Yan Wang, Humphrey O Obie, Zhuying Li, Flora D Salim, John Grundy, and Florian 'Floyd' Mueller. 2024. GustosonicSense: Towards understanding the design of playful gustosonic eating experiences. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–12.

[171] Yong Wang, Tianyu Yang, Chunxiao Wang, Feng Li, Pengfei Hu, and Yiran Shen. 2024. BudsAuth: Towards Gesture-Wise Continuous User Authentication Through Earbuds Vibration Sensing. *IEEE Internet of Things Journal* (2024).

[172] Yuntao Wang, Xiyuxing Zhang, Jay M Chakalasiya, Xuhai Xu, Yu Jiang, Yuang Li, Shwetak Patel, and Yuanchun Shi. 2022. HearCough: Enabling continuous cough event detection on edge computing hearables. *Methods* 205 (2022), 53–62.

[173] 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 (2022), 1–24.

[174] Zi Wang, Yilin Wang, and Jie Yang. 2024. Earslide: a secure ear wearables biometric authentication based on acoustic fingerprint. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–29.

[175] Zeyu Wang, Ruotong Yu, Xiangyang Wang, Jiexin Ding, Jiankai Tang, Jun Fang, Zhe He, Zhuojun Li, Tobias Röddiger, Weiye Xu, et al. 2025. Computing with smart rings: A systematic literature review. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–54.

[176] Jordan Waters, Jake Stuchbury-Wass, Yang Liu, Kayla-Jade Butkow, and Cecilia Mascolo. 2025. Deep-learning based segmentation of in-ear cardiac sounds. In *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1–7.

[177] Wentao Xie, Qingyong Hu, Jin Zhang, and Qian Zhang. 2023. Earspiro: Earphone-based spirometry for lung function assessment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–27.

[178] 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 *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 1789–1798.

[179] Chi Xu, Wentao Xie, Baichen Yang, Yizhen Zhang, Yanbin Gong, Jin Zhang, Wei Li, Shifang Yang, and Qian Zhang. 2025. EasySpiro: Assessing Lung Function via Arbitrary Exhalations on Commodity Earphones. (2025).

[180] Vasco Xu, Chenfeng Gao, Henry Hoffmann, and Karan Ahuja. 2024. MobilePoser: Real-Time Full-Body Pose Estimation and 3D Human Translation from IMUs in Mobile Consumer Devices. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–11.

[181] Yuchen Xu, Ernesto De la Paz, Akshay Paul, Kuldeep Mahato, Juliane R Sempionatto, Nicholas Tostado, Min Lee, Gopabandhu Hota, Muyang Lin, Abhinav Uppal, et al. 2023. In-ear integrated sensor array for the continuous monitoring of brain activity and of lactate in sweat. *Nature Biomedical Engineering* 7, 10 (2023), 1307–1320.

[182] Zhenye Xu and Yao Guo. 2024. GCCRR: A Short Sequence Gait Cycle Segmentation Method Based on Ear-Worn IMU. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 650–654.

[183] Qiuyue Xue, Yujia Liu, Vikram Iyer, and Shwetak Patel. 2024. Design of Thermal Earring: A Low-Power Wireless Earring for Longitudinal Temperature Monitoring. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 366–370.

[184] Qiuyue Xue, Dilini Nissanka, Jiachen Tammy Yan, Ruiqing Wang, Shwetak Patel, and Vikram Iyer. 2025. PPG Earring: Wireless Smart Earring for Heart Health Monitoring. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–16.

[185] Qiang Yang, Yang Liu, Jake Stuchbury-Wass, Kayla-Jade Butkow, Dong Ma, and Cecilia Mascolo. 2024. BrushBuds: Toothbrushing Tracking Using Earphone IMUs. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 655–660.

[186] Qiang Yang, Yang Liu, Jake Stuchbury-Wass, Kayla-Jade Butkow, and Cecilia Mascolo. 2025. SmarTeeth: Augmenting Manual Toothbrushing with In-ear Microphones. (2025).

[187] Qiang Yang, Yang Liu, Jake Stuchbury-Wass, Mathias Ciliberto, Tobias Röddiger, Kayla-Jade Butkow, Adam Luke Pullin, Emeli Panariti, Dong Ma, and Cecilia Mascolo. 2025. HearForce: Force Estimation for Manual Toothbrushing with Earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–22.

[188] Qiang Yang, Jake Stuchbury-Wass, Thomas A Bytheway, and Cecilia Mascolo. 2025. OpenBrush: An Open-Source Manual Toothbrush for Force and Motion Sensing. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1297–1300.

[189] Qiang Yang and Yuanqing Zheng. 2022. Deepear: Sound localization with binaural microphones. *IEEE Transactions on Mobile Computing* 23, 1 (2022), 359–375.

[190] Yongjie Yang, Tao Chen, Zhenlin An, Shirui Cao, Xiaoran Fan, and Longfei Shangguan. 2025. LeakyFeeder: In-Air Gesture Control Through Leaky Acoustic Waves. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*. 144–157.

[191] Yongjie Yang, Tao Chen, Yujing Huang, Xiuzhen Guo, and Longfei Shangguan. 2024. MAF: Exploring mobile acoustic field for hand-to-face gesture interactions. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–20.

[192] Xianrong Yao, Chengzhang Yu, Lingde Hu, Yincheng Jin, Yang Gao, and Zhanpeng Jin. 2025. IMUFace: Real-Time, Low-Power, Continuous 3D Facial Reconstruction Through Earphones. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*. 614–615.

[193] Masaki Yasuhara, Isao Nambu, and Shohei Yano. 2022. Bilateral ear acoustic authentication: A biometric authentication system using both ears and a special earphone. *Applied Sciences* 12, 6 (2022), 3167.

[194] Xing Yi, Samith S Herath, Hiroshi Ogawa, Hiroki Kuroda, Kosuke Oiwa, and Shusaku Nomura. 2024. Human pulse wave detection with consumer earphones and headphones. *Artificial Life and Robotics* 29, 1 (2024), 22–28.

[195] Taiki Yuma and Kazuya Murao. 2025. Mouth Gesture Recognition Using PPG Sensors in Earbuds. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1024–1030.

[196] Linghan Zhang, Sheng Tan, Zi Wang, Yili Ren, Zhi Wang, and Jie Yang. 2020. Viblive: A continuous liveness detection for secure voice user interface in iot environment. In *Proceedings of the 36th Annual Computer Security Applications Conference*. 884–896.

[197] Ruidong Zhang, Hao Chen, Devansh Agarwal, Richard Jin, Ke Li, François Guimbretière, and Cheng Zhang. 2023. Hpspeech: Silent speech interface for commodity headphones. In *Proceedings of the 2023 ACM International Symposium on Wearable Computers*. 60–65.

[198] Shijia Zhang, Taiting Lu, Hao Zhou, Yilin Liu, Runze Liu, and Mahanth Gowda. 2023. I am an earphone and i can hear my user's face: Facial landmark tracking using smart earphones. *ACM Transactions on Internet of Things* 5, 1 (2023), 1–29.

[199] Shibo Zhang, Ebrahim Nemati, Minh Dinh, Nathan Folkman, Tousif Ahmed, Mahbubur Rahman, Jilong Kuang, Nabil Alshurafa, and Alex Gao. 2022. Coughtrigger: Earbuds imu based cough detection activator using an energy-efficient sensitivity-prioritized time series classifier. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[200] Xiyuxing Zhang, Yuntao Wang, Yuxuan Han, Chen Liang, Ishan Chatterjee, Jiankai Tang, Xin Yi, Shwetak Patel, and Yuanchun Shi. 2024. The earsavas dataset: Enabling subject-aware vocal activity sensing on earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (2024), 1–26.

[201] Xiyuxing Zhang, Yuntao Wang, Jingru Zhang, Yaqing Yang, Shwetak Patel, and Yuanchun Shi. 2023. Earcough: Enabling continuous subject cough event detection on hearables. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–6.

[202] Guangrong Zhao, Yiran Shen, Feng Li, Lei Liu, Lizhen Cui, and Hongkai Wen. 2024. Ui-Ear: On-face Gesture Recognition Through On-ear Vibration Sensing. *IEEE Transactions on Mobile Computing* (2024).

[203] Yankai Zhao, Jin Zhang, Jiao Li, and Tao Sun. 2025. PalateTouch: Enabling Palate as a Touchpad to Interact with Earphones Using Acoustic Sensing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–22.

[204] Zhiyuan Zhao, Fan Li, Yadong Xie, Huanran Xie, Kerui Zhang, Li Zhang, and Yu Wang. 2024. HearBP: Hear Your Blood Pressure via In-ear Acoustic Sensing Based on Heart Sounds. In *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*. IEEE, 991–1000.

[205] Hao Zhou, Md Mahbubur Rahman, Mehrab Bin Morshed, Yunzhi Li, Md Saiful Islam, Larry Zhang, Jungmok Bae, Christina Rosa, Wendy Berry Mendes, and Jilong Kuang. 2025. Know Your Heart Better: Multimodal Cardiac Output Monitoring using Earbuds. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.

[206] Yuzheng Zhu, Zhangxin Liang, Jie Zheng, Yongpan Zou, Victor CM Leung, and Kaishun Wu. 2025. VitalEar: an Earable Heartbeat and Respiratory Rate Monitoring System under Aerobic Exercises. *IEEE Transactions on Mobile Computing* (2025).