Earables: Wearable Computing on the Ears

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

von der KIT-Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

M.Sc. Tobias Röddiger

Tag der mündlichen Prüfung: 26. Juli 2023

1. Referent: Prof. Dr. Michael Beigl

2. Referent: Prof. Dr. Hans Gellersen

Acknowledgements

I would like to express my deepest gratitude to my primary advisor, Prof. Dr. Michael Beigl, for his unlimited support and advice to guide me through my dissertation. Thanks for always listening to my ideas, trusting in my skills, and putting me into the positions where I could grow and learn the most, to become the researcher I am today. I could not have imagined having a better advisor and mentor for my PhD.

Also, thank you to my second advisor, Prof. Dr. Hans Gellersen (Lancaster University), for being a steady companion and trustworthy source of advice for many years. I would also like to thank Prof. Dr. Joseph A. Paradiso (MIT Media Lab) for giving me the opportunity to work with his group and hosting me at MIT. Also, thank you to PD Dr. Victor Pankratius (Bosch Sensortec GmbH) for the exciting industry insights and project challenges. A special thanks to Dr. hc. Hans-Werner Hector and Josephine Hector, whose support has inspired me to pursue a career in computer science in the first place.

I am very grateful to have met Dr. Christopher Clarke (University of Bath). Not only was he an incredibly smart colleague to work with, but more importantly, he also has become a great friend of mine along the way.

Thanks to my wonderful colleagues at KIT - Dr. Anja Exler, Dr. Erik Pescara, Haibin Zhao, Dr. Matthias Budde, Dr. Michael Hefenbrock, Paula Breitling, Dr. Till Riedel, Tim Schneegans, Tobias King, and Yexu Zhou. A special thanks to Fangzheng Liu from MIT for letting me participate in the AstroAnt project, and Michael Knierim from KIT for collaborating on the ExG headphones. Thanks to Denise Hillmann, Helga Scherer, Melissa Alpman, and Zina Tsiouma for the support with all administrative tasks.

A massive thanks to all the students that I have supervised or worked with. Your enthusiasm to learn and drive have many times motivated me to exceed my own expectations. It was a pleasure working with my research assistants

Anja Hansen, Daniel Wolffram, Dennis Moschina, Dylan Ray Roodt, Jennifer Münk, Leon Simmon, Mark Schenkel, Michael Küttner, Murat Kurnaz, Nils Kerwer, Oliver Bagge, Ömer Erdinç Yağmurlu, Philipp Lepold, Shanshan Du, Titus von Plüskow, and Vladimir Bashkuev.

Thank you Laura Traub for being a loving, intelligent, and ever-supporting girlfriend. Thanks for giving me the freedom and strength to spend countless days and nights on pushing my research journey forward.

I want express my deepest gratitude to my grandma, Brigitte Seider, for teaching me the value of love, kindness, and happiness. Also, I am eternally grateful to my grandpa, Max Heinz Seider, for always supporting me and challenging me many times to better understand how the world functions. Thanks to my grandmother Hildegard Röddiger for teaching me the importance of humbleness and gratefulness.

Thank you to my sister Julia Röddiger for being a never ending source of cheerfulness, and inspiration for creativity and design.

And most importantly, thanks to my lovely parents Angela Seider-Röddiger and Ralf Röddiger, for always supporting me in every possible way and trusting me with becoming who I am today. All the things you have taught me and the values you passed on to me are fundamental to all my achievements. I owe you more than I could ever express and I dedicate this dissertation to you.

Stay healthy!

Abstract

Earphones have gained widespread consumer adoption due to their ability to provide private audio channels for listening to music, watching the latest movies during commuting, or making hands-free phone calls. Because of this clear primary purpose, earphones have seen broader adoption than other wearables, such as smart glasses. In recent years, a new class of wearable computing devices known as "earables" has emerged. These devices are designed to be worn in or around the ears and integrate various sensors to enhance the functionalities of earphones. The unique proximity of earables to key anatomical structures of the human body offers a distinct platform for sensing a diverse range of properties, processes, and activities.

While some progress has been made in the realm of earables research, vast potential remains unexplored. Thus, the purpose of this dissertation is to provide novel insights into the capabilities of earables by introducing advanced sensing approaches that enable the detection of previously inaccessible phenomena. Through the introduction of new state-of-the-art hardware and algorithms, this dissertation aims to push the boundaries of what is achievable in the domain of earables, ultimately striving to guide them towards a general-purpose sensing platform for human augmentation.

To establish a robust foundation, this dissertation systematically synthesizes the state-of-the-art in earable sensing works and presents a uniquely comprehensive taxonomy of earable sensing capabilities based on 271 relevant publications. By connecting low-level sensing principles with higher-level phenomena, the dissertation then goes on to summarize works across various domains, including (i) physiological monitoring and health, (ii) movement and activity, (iii) interaction, and (iv) authentication and identification.

Expanding upon existing research in the field of physiological monitoring and health using earables, this dissertation introduces algorithms, statistical

evaluations, and empirical evidence to demonstrate the feasibility of tracking respiration rates and detecting episodes of increased coughing frequency through the utilization of in-ear accelerometers and gyroscopes. These novel sensing capabilities offer new ways for earables to promote healthier lifestyles and enable proactive healthcare.

Furthermore, this dissertation introduces an eye-tracking approach termed "earEOG" to facilitate activity tracking. By systematically evaluating electrode potentials measured around the ears in a headphone form factor, the dissertation unveils a new avenue for measuring gaze in a less intrusive and more comfortable manner. The dissertation compares different electrode positions and introduces a regression model to predict absolute gaze angles from earEOG. This development opens up new opportunities for research that seamlessly integrates into daily life, providing deeper insights into human behavior and also shows how the unique form factors of earables, combined with suitable sensing, enables the detection of new phenomena.

To advance the interaction possibilities of earables, this dissertation introduces a discreet input technique called "EarRumble", which relies on voluntary control of the tensor tympani muscle in the middle ear. The dissertation offers insights into the prevalence, ease of use, and comfort of EarRumble, along with practical applications in two real-world scenarios. The EarRumble approach effectively transforms the ear from a purely receptive organ into a dual function one, enabling it not only to receive, but also to generate output signals. It essentially introduces the ears as an additional interactive medium, bridging the gap between humans and machines in a hands-free and eyes-free manner. This contribution unveils an interaction technique that users describe as "magical and almost telepathic", highlighting significant untapped potential within the realm of earables.

Finally, the findings across the different application domains and research insights accumulate in an open-source hard- and software earable sensing platform called "OpenEarable". OpenEarable encompasses a range of advanced sensing capabilities suitable for various earable research applications, while also being easy to manufacture. This reduces entry barriers into ear-based sensing research and, therefore, OpenEarable aids in unleashing the full potential that earables may offer. In addition, the dissertation contributes fundamental design guidelines and reference architectures for earables. Through this research, the dissertation bridges the gap between foundational research on earable sensing capabilities and their practical deployment in real-world scenarios.

In sum, the dissertation contributes new usage scenarios, algorithms, hardware prototypes, statistical evaluations, empirical evidence, and design guidelines to advance the field of earable computing. Furthermore, this dissertation extends the traditional scope of earphones, transitioning them from audio focused devices towards a platform offering a plethora of advanced sensing capabilities to understand properties, processes, and activities. This redefinition puts earables at the frontier of becoming a significant wearable category so that the vision of earables to be a general-purpose sensing platform for human augmentation becomes increasingly apparent.

Zusammenfassung

Kopfhörer haben sich bei Verbrauchern durchgesetzt, da sie private Audiokanäle anbieten, zum Beispiel zum Hören von Musik, zum Anschauen der neuesten Filme während dem Pendeln oder zum freihändigen Telefonieren. Dank diesem eindeutigen primären Einsatzzweck haben sich Kopfhörer im Vergleich zu anderen Wearables, wie zum Beispiel Smartglasses, bereits stärker durchgesetzt. In den letzten Jahren hat sich eine neue Klasse von Wearables herausgebildet, die als "Earables" bezeichnet werden. Diese Geräte sind so konzipiert, dass sie in oder um die Ohren getragen werden können. Sie enthalten verschiedene Sensoren, um die Funktionalität von Kopfhörern zu erweitern. Die räumliche Nähe von Earables zu wichtigen anatomischen Strukturen des menschlichen Körpers bietet eine ausgezeichnete Plattform für die Erfassung einer Vielzahl von Eigenschaften, Prozessen und Aktivitäten.

Auch wenn im Bereich der Earables-Forschung bereits einige Fortschritte erzielt wurden, wird deren Potenzial aktuell nicht vollständig abgeschöpft. Ziel dieser Dissertation ist es daher, neue Einblicke in die Möglichkeiten von Earables zu geben, indem fortschrittliche Sensorikansätze erforscht werden, welche die Erkennung von bisher unzugänglichen Phänomenen ermöglichen. Durch die Einführung von neuartiger Hardware und Algorithmik zielt diese Dissertation darauf ab, die Grenzen des Erreichbaren im Bereich Earables zu verschieben und diese letztlich als vielseitige Sensorplattform zur Erweiterung menschlicher Fähigkeiten zu etablieren.

Um eine fundierte Grundlage für die Dissertation zu schaffen, synthetisiert die vorliegende Arbeit den Stand der Technik im Bereich der ohr-basierten Sensorik und stellt eine einzigartig umfassende Taxonomie auf der Basis von 271 relevanten Publikationen vor. Durch die Verbindung von Low-Level-Sensor-Prinzipien mit Higher-Level-Phänomenen werden in der Dissertation anschließend Arbeiten aus verschiedenen Bereichen zusammengefasst, darunter (i) phys-

iologische Überwachung und Gesundheit, (ii) Bewegung und Aktivität, (iii) Interaktion und (iv) Authentifizierung und Identifizierung.

Diese Dissertation baut auf der bestehenden Forschung im Bereich der physiologischen Überwachung und Gesundheit mit Hilfe von Earables auf und stellt fortschrittliche Algorithmen, statistische Auswertungen und empirische Studien vor, um die Machbarkeit der Messung der Atemfrequenz und der Erkennung von Episoden erhöhter Hustenfrequenz durch den Einsatz von In-Ear-Beschleunigungsmessern und Gyroskopen zu demonstrieren. Diese neuartigen Sensorfunktionen unterstreichen das Potenzial von Earables, einen gesünderen Lebensstil zu fördern und eine proaktive Gesundheitsversorgung zu ermöglichen.

Darüber hinaus wird in dieser Dissertation ein innovativer Eye-Tracking-Ansatz namens "earEOG" vorgestellt, welcher Aktivitätserkennung erleichtern soll. Durch die systematische Auswertung von Elektrodenpotentialen, die um die Ohren herum mittels eines modifizierten Kopfhörers gemessen werden, eröffnet diese Dissertation einen neuen Weg zur Messung der Blickrichtung. Dabei ist das Verfahren weniger aufdringlich und komfortabler als bisherige Ansätze. Darüber hinaus wird ein Regressionsmodell eingeführt, um absolute Änderungen des Blickwinkels auf der Grundlage von earEOG vorherzusagen. Diese Entwicklung eröffnet neue Möglichkeiten für Forschung, welche sich nahtlos in das tägliche Leben integrieren lässt und tiefere Einblicke in das menschliche Verhalten ermöglicht. Weiterhin zeigt diese Arbeit, wie sich die einzigarte Bauform von Earables mit Sensorik kombinieren lässt, um neuartige Phänomene zu erkennen.

Um die Interaktionsmöglichkeiten von Earables zu verbessern, wird in der vorliegenden Dissertation eine diskrete Eingabetechnik namens "EarRumble" vorgestellt, die auf der freiwilligen Kontrolle des Tensor Tympani Muskels im Mittelohr beruht. Die Dissertation bietet Einblicke in die Verbreitung, die Benutzerfreundlichkeit und den Komfort von EarRumble, zusammen mit praktischen Anwendungen in zwei realen Szenarien. Der EarRumble-Ansatz erweitert das Ohr von einem rein rezeptiven Organ zu einem Organ, das nicht nur Signale empfangen, sondern auch Ausgangssignale erzeugen kann. Im Wesentlichen wird das Ohr als zusätzliches interaktives Medium eingesetzt, welches eine freihändige und augenfreie Kommunikation zwischen Mensch und Maschine ermöglicht. EarRumble stellt eine Interaktionstechnik vor, die von den Nutzern als “magisch und fast telepathisch” beschrieben wird, und zeigt ein erhebliches ungenutztes Potenzial im Bereich der Earables auf.

Aufbauend auf den vorhergehenden Ergebnissen der verschiedenen Anwendungsbereiche und Forschungserkenntnisse mündet die Dissertation in einer offenen Hard- und Software-Plattform für Earables namens “OpenEarable”. OpenEarable umfasst eine Reihe fortschrittlicher Sensorfunktionen, die für verschiedene ohrbasierte Forschungsanwendungen geeignet sind, und ist gleichzeitig einfach herzustellen. Hierdurch werden die Einstiegshürden in die ohrbasierte Sensorforschung gesenkt und OpenEarable trägt somit dazu bei, das gesamte Potenzial von Earables auszuschöpfen. Darüber hinaus trägt die Dissertation grundlegenden Designrichtlinien und Referenzarchitekturen für Earables bei. Durch diese Forschung schließt die Dissertation die Lücke zwischen der Grundlagenforschung zu ohrbasierten Sensoren und deren praktischem Einsatz in realen Szenarien.

Zusammenfassend liefert die Dissertation neue Nutzungsszenarien, Algorithmen, Hardware-Prototypen, statistische Auswertungen, empirische Studien und Designrichtlinien, um das Feld des Earable Computing voranzutreiben. Darüber hinaus erweitert diese Dissertation den traditionellen Anwendungsbereich von Kopfhörern, indem sie die auf Audio fokussierten Geräte zu einer Plattform erweitert, welche eine Vielzahl fortschrittlicher Sensorfähigkeiten bietet, um Eigenschaften, Prozesse und Aktivitäten zu erfassen. Diese Neuausrichtung ermöglicht es Earables sich als bedeutende Wearable Kategorie zu etablieren, und die Vision von Earables als eine vielseitige Sensorenplattform zur Erweiterung der menschlichen Fähigkeiten wird somit zunehmend realer.

Contents

List of Figures

List of Tables

1. Introduction

Earables are wearable computing devices that are worn in or around the ear. Ear-worn devices evolved for isolated purposes, as hearing aids and personal speakers. In using the notion of "earables", this dissertation refers to devices that integrate wider capabilities, as a new type of ubiquitous computing platform. In consumer electronics, earphones have already become wireless at large scale, and increasingly integrate diverse types of sensors to extend their functionality [397]. Conversely, hearing aids integrate sensing to personalise sound amplification but also converge toward wireless integration with other devices [255]. These trends are mirrored in research, where earables (and synonymously "hearables") have emerged as a distinct area of investigation [379; 235; 102]. At the heart of much of the research in this new field are questions of sensing - what can be detected and observed with earables, and what interactions and applications are enabled by sensing in or on the ear?

To this extent, this dissertation seeks to: (i) synthesize the state-of-the-art in earable sensing works; and (ii) expand upon the status quo through the contributions of novel sensing capabilities of earables, and evaluations thereof.

1.1 Motivation

Earables, with their specific positioning on the human body, provide a distinct platform for sensing of a wide range of properties, processes and activities. They are portable and their small and lightweight form factor allows them to be worn for prolonged periods throughout the day. The shape of the ear affords a variety of mechanical anchoring points [184; 236; 359] and the ears are less susceptible to motion disturbance and artefacts as the body stabilises the head during locomotion [164; 232]. The proximity to the brain and blood vessels enables the accurate measurement of brain activity, cyclic blood flow and related properties [137], and the inner ear cavity acts as an echo chamber to amplify internal body sounds [13]. The position of the ears on the head allows for a multitude of facial, neck, and eye muscle activations to be sensed [16] and input from head movement [16], facial gestures [304], mouth movements [435], and eye gaze to be detected [374; 57]. The ear itself is easily and comfortably reached by the hands [494; 241], while the distinctive surface area creates opportunities for a variety of touch interactions [272]. In sum, earables are capable of sensing a wide variety of processes of the skeletal (e.g., gait [32]), muscular (e.g., facial expressions [304]), nervous (e.g., brain activity [115]), endocrine (e.g., emotions [35]), cardiovascular (e.g., blood pressure [76]), respiratory (e.g., breathing [450]), and digestive (e.g., food intake [150]) systems).

1.2 Aim

Earables have the potential to be a ubiquitous platform, but the understanding of their capabilities and limitations is still evolving. Despite the extensive research conducted on earables, achieving reliable and consistent sensing results remains a challenge, with much untapped potential yet to be explored.

This dissertation aims to provide fundamentally new insights into the capabilities of earables, contributing novel sensing approaches that enable the detection of phenomena previously inaccessible to earable devices. By breaking ground with the introduction of new hardware and algorithms, this dissertation seeks to push the boundaries of what is possible in the earables domain, with the ultimate goal to steer earables towards a general-purpose sensing platform.

1.3 Objectives

To achieve the overarching aim, the dissertation focuses on the development of solutions that overcome current limitations of earable sensing capabilities. On a conceptual level, this leads to the following objectives:

1. Thoroughly understanding the existing challenges and limitations in earable sensing technologies through systematic literature review.

2. Identifying novel sensing opportunities and under-explored phenomena

in the earables domain.

3. Designing and implementing innovative hardware and algorithms tailored to these novel sensing opportunities through prototyping.

4. Evaluating the effectiveness of the proposed solutions in terms of reliability and usability through empirical studies.

5. Integrating the proposed solutions into a unified, general-purpose earable sensing platform.

Through this approach, this dissertation seeks to facilitate a future where earables become a ubiquitous, integral part of our lives, serving as a general-purpose sensing platform.

1.4 Contributions

In line with the overarching aim and objectives, this dissertation presents a series of contributions that guide earables towards the goal of becoming a general-purpose sensing platform.

1.4.1 Taxonomy of Sensing on the Ears

To establish a foundation for exploring novel ear-based sensing applications, the first objective of this dissertation is to comprehensively understand the existing research on earables, to reveal what has been accomplished so far and how. This leads to the first research question of this dissertation:

> **Question 1:** What can be detected and observed with earables, and what interactions and applications are enabled by sensing in or on the ears?

To address this question, a systematic literature review of 271 peer-reviewed research articles was conducted. Each article reviewed was classified with respect to sensing principles applied, types of information gained, and purposes to which sensing was used. Through iteration, this process resulted in the development of a taxonomy of *phenomena* sensed. At the lowest level, the dissertation identifies and characterises phenomena that are directly sensed with sensors placed in or on the ear as *fundamental phenomena*, including for instance motion or blood perfusion. Other phenomena are identified as indirectly observable and derived from fundamental phenomena, ranging from physiological parameters (e.g., heart rate) and lower-level cues (e.g., earable state) to conditions (e.g., stress), actions (e.g., gestures), activities (e.g., daily tasks) and other context (e.g., identity). The dissertation shows how higher-level phenomena build on fundamental phenomena, and relates this to different sensors that have been investigated for their observation. The result is an open-ended taxonomy that provides a clear end-to-end structure for classification of earable sensing work, from sensors employed to fundamental and higher-level phenomena and their application. This results in the following contributions:

> **Contribution 1:** (1) A uniquely comprehensive, open-ended taxonomy of sensing on the ears, identifying 13 fundamental phenomena using 21 different sensors to enable close to 50 phenomena to be sensed around the ears; (2) a summary of related works in four main areas (i) physiological monitoring and health, (ii) movement and activity, (iii) interaction, and (iv) authentication and identification.

1.4.2 Measuring Respiration and Cough with Inertial Sensing

From the fundamental understanding for earable sensing, it was quickly found that the unique positioning of earables has led to research on their potential for health and physiological monitoring, including tracking vital signs and detecting medical conditions. Still, the potential of earables for health monitoring has not been fully realized. Accelerometers and gyroscopes, which measure health- and fitness-related body motions, represent a promising set of sensors. They are cost-effective, energy-efficient, and already integrated into commodity earables (e.g., Apple AirPods [20]). Therefore, the second objective of this dissertation is to explore the use of these sensors for new health applications. This leads to the following research question:

> **Question 2:** How can inertial sensing on the ears be applied to realize respiration rate sensing while the user is at rest?

To address the question, breath data in three different poses (sitting, standing, supine) was collected from 12 participants wearing a 6-axis inertial measurement unit on the left ear (accelerometer and gyroscope). They were instructed to breath naturally before and after exercising. Based on this data, a filter-based processing pipeline and the necessary tuning parameters were introduced to achieve an absolute error of 2.62 and 2.55 breath cycles per minute for accelerometer and gyroscope, respectivley. This performance is close to the accuracy required for non-vital signs monitoring. Hence, the dissertation contributes:

> **Contribution 2:** Algorithms and empirical evidence to show the feasibility of respiration rate tracking based on in-ear accelerometer and gyroscope data in three stationary poses (sitting, standing, supine).

Breathing is a critical vital sign that can be an early warning indicator of physiological deterioration [363]. Especially with the relevance of COVID-19 during the writing period of this dissertation, respiratory illness was a particularly prominent topic. To contain a pandemic, detecting illness early on can help isolating patients which helps containing a virus outbreak [217]. Hence detecting cough events with accelerometer and gyroscope data with earables appeared to be a promising path for further investigation which lead to the following research question posed by this dissertation:

> **Question 3:** How can inertial sensing on the ears be applied for detecting episodes of increased cough frequency (e.g., during illness)?

To answer this question, the dissertation evaluates 4,200 throat activity related event samples to propose a machine learning model for detecting cough events and identifying episodes of increased cough based on statistical analysis. The classifier achieves 77% accuracy in predicting cough events. Based on the developed machine learning model, the dissertation presents a statistical approach to discriminate episodes of increased cough after a defined observation period (e.g., 24 hours). Therefore, the dissertation contributes:

> **Contribution 3:** Algorithms, empirical evidence, and statistical modelling to realize episode of increased cough detection under motion stress (e.g., walking) based on in-ear accelerometer and gyroscope data.

1.4.3 Eye Tracking based on Periauricular Electric Potentials

Activity tracking is a popular domain in ubiquitous computing. An interesting modality to track the activity of a person is eye tracking based on electric potentials of the eyes [82]. While it is well-known that movement of the eyes can be sensed with electrodes around the ears, no efforts were undertaken to understand this phenomenon systematically. This leads to the following question:

> **Question 4:** How well can electric potentials sensed around the ears be leveraged as an alternative for gold-standard eye tracking and where to position measurement electrodes to yield the best results?

Hence, the third contribution of this dissertation explores ear Electrooculography (earEOG) using 28 electrodes positioned around the ears in a regular headphone form factor (14 on earch ear). Based on a data collection study with 3 participants, the dissertation identifies the most effective electrode pairs and establishes their correlation to gold-standard EOG with respect to different eye movement patterns. In the four cardinal directions (left, right, up, down) a regression model based on earEOG achieves an average absolute angular error of $9.2° \pm 4.5°$ to predict saccades between 5° and 30° (vs. $4.4° \pm 1.1°$ gold-standard EOG). In sum, the contributions are:

> **Contribution 4:** (1) An investigation of EOG-based eye tracking using electrodes placed around the ears in a custom-built headphone form factor, providing earEOG - a novel approach to on-the-go wearable eye tracking; (2) an evaluation of different electrode positions for earEOG and recommendations for the optimal placement of electrodes, to enable a more effective use of earEOG for eye tracking; and (3) an evaluation of earEOG to predict absolute gaze angles of the four cardinal directions.

1.4.4 Discreet, Hands- & Eyes-Free Tensor Tympani Input

While the former research questions focus on passively sensing physiological processes and activities, a lot of research has also been conducted on how user can actively perform input into ear-worn devices. Current interaction techniques with earables mainly relied on visible gestures or vocal commands. By exploring new ways to interact with earables that do not require the use physically demanding gestures or voice, earables could be more accessible for a wider range of users, including those with disabilities or who work in environments where voice commands or gestures are not practical or desirable. Therefore, this dissertation explores how, compared to existing principles, input may be offered in a discreet as well as hands- and eyes-free manner. This leads to the fifth research question of this dissertation:

> **Question 5:** How may the tensor tympani muscle, found in the human middle ear, be leveraged for interaction with earables as a discreet, hands- and eyes-free technique?

To answer this question, the dissertation introduces a novel interaction technique based on the tensor tympani muscle, called "ear rumbling". This technique is based on the voluntary contraction of the tensor tympani. Through an online questionnaire, the dissertation sheds light on the prevalence of the voluntary tensor tympani contraction ability, with 43.2% of participants self-reporting the controlling ability (N = 192). Using a custom-built sensor earpiece, the thesis also introduces a sensing technique to detect rumbles via ear canal pressure and describes the required processing pipeline to use ear rumbling as a real-time interaction technique. The developed system yields 95% accuracy in detecting three ear rumble gestures (N = 16). Furthermore, the dissertation characterizes ear rumbling with respect to reaction time. In a task-based study with eight users ear rumbling is applied to interact with a music player and handle phone calls. Users described the technique as "magical" and "almost telepathic". Through these studies, the dissertation does the following contributions:

> **Contribution 5:** (1) EarRumble, a hands- and eyes-free, discreet input technique based on voluntary control of the tensor tympani muscle and sensed through in-ear barometry; (2) an indication of how prevalent ear rumbling is, how easy it is to perform on demand, and how comfortable it is through an online questionnaire; (3) data-driven insights into how well users can contract the tensor tympani muscle in the context of interaction; and (4) insights into how EarRumble can be used for interaction, grounded in real-world applications.

1.4.5 Advancing towards General Purpose Ear-based Sensing

Earables have been extensively studied due to their sensing capabilities and potential applications. However, many researchers often build custom devices from scratch which requires substantial effort and makes it more challenging to reproduce results. Therefore, this dissertation seeks to understand:

> **Question 6:** How can earables be established as a general-purpose platform for prototyping sensor-based applications focused around the ears?

Therefore, and by taking all previous research into account, this dissertation introduces a new open-source hardware platform and accompanying software toolchain for earable sensing research. The platform is called OpenEarable and has a 9-axis IMU, an ear canal ultrasound microphone and speaker, and an ear canal pressure and temperature sensor. OpenEarable is implemented as an earhook design and the earpiece inside the ear canal can be easily exchanged to allow for other sensors to be applied on the ears. The dissertation shows the wide capabilities of OpenEarable based on three sample applications: detecting jaw motions, tracking body movement, and measuring ear canal sound reflections for authentication. Moreover, based on 39 student projects, this dissertation demonstrates how such earable prototyping platforms can be used for problem-based learning and generating new ideas for earable research. The introduction of OpenEarable will hopefully reduce the assessment heterogeneity of studies conducted with earables and contribute to the establishment of earables as a general-purpose platform.

Therefore, this dissertation seeks to establish a general-purpose earable prototyping platform resulting in the following contributions:

> **Contribution 6:** (1) An open-source hard- and software platform for earable sensing applications called "OpenEarable" which features a 9-axis inertial measurement unit, an ear canal pressure and temperature sensor, as well as an inward facing ultrasound microphone and speaker; (2) an evaluation of earable prototyping platforms to show that such devices enable new ideas and create interest in the research topic.

While earables are commonly used during the day, their broad sensing capabilities are also interesting to track sleep-related parameters (e.g., brain activity [374]). To understand comfort of earables during sleep and establish best practices, the dissertation explores the wearability of seven different earable form factors for sleep. In a 14 participant study, users wore each of the devices for one night. It was found that rigid parts should be placed behind the ears and obstruction of the ear canal should be avoided. All earables had an adverse effect on sleep quality. This results in the following contribution:

> **Contribution 7:** An evaluation of the comfort and wearability of seven different earable form factors during sleep

1.5 Structure

The dissertation is divided into eight logical sections that delve into specific aspects of earable technology and its applications. The structure of the thesis is represented as a matrix in Figure 1.1, which provides an overview of the relationships between the different sections and subsections. The matrix format allows the reader to quickly understand how the different sections of the dissertation relate to each other and how they fit together as a whole. Overall, the structure of the dissertation is designed to provide a comprehensive exploration of earable technology and its potential applications in different domains.

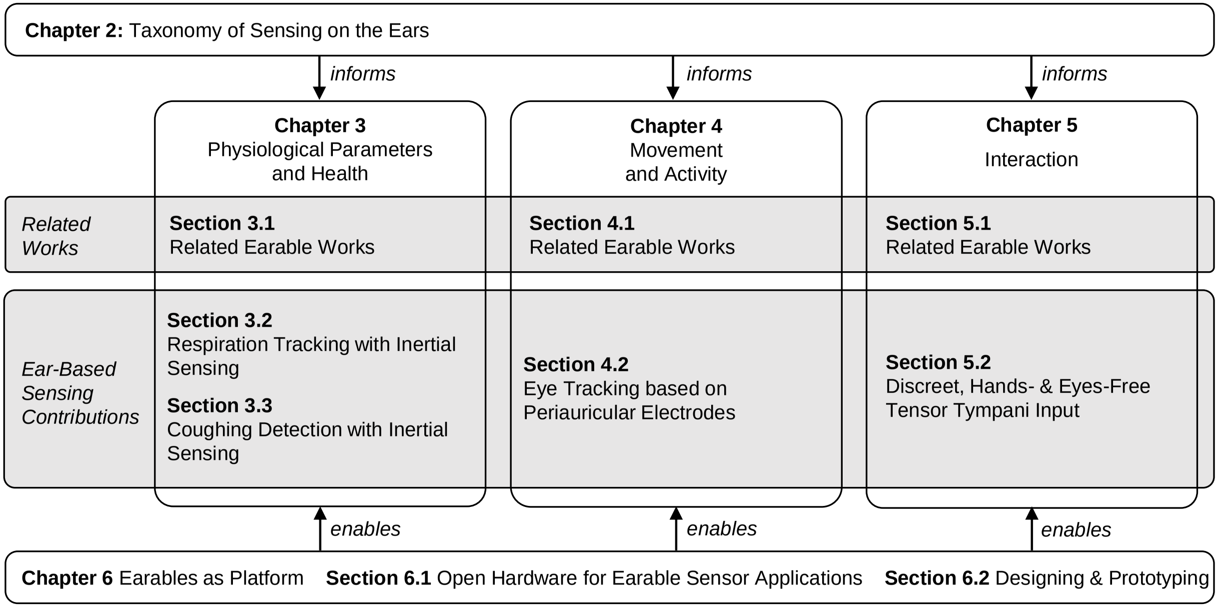


Figure 1.1.: Matrix relationship between the different sections covering different aspects of the earable research field. Chapter 1 (Introduction), Chapter 7 (Discussion) and Chapter 8 (Conclusion) are excluded from the figure because they do not pertain to specific aspects of the earable research field, but rather provide context, analysis, and overall conclusions.

After initially establishing the foundation for ear-based sensing based on the taxonomy presented in chapter 2, the dissertation follows the primary research areas identified by the literature review: (i) physiological parameters and health in chapter 3, (ii) movement and activity in chapter 4, and (iii) interaction in chapter 5. The dissertation presents related works across these domains and introduces the application-specific ear-based sensing contributions: respiration rate sensing (section 3.2), episode of increased cough detection (section 3.3), eye tracking with periauricular electric potentials (section 4.2), and discreet hands- and eyes-free tensor tympani muscle input (section 5.2). Based on all insights generated, chapter 6 then investigates earables as a platform with respect to prototyping and design. This includes the introduction of the OpenEarable hardware and software platform for ear-based sensor applications.

The application-specific chapters are followed by discussions of the earable field as a whole in chapter 7, focusing on research challenges and gaps that remain open and to be investigated for future research. In addition, the discussion includes a summary of related works of the fourth research field identified in this dissertation but not investigated through specific applications: (iv) authentication and identification (section 7.4). In the last chapter 8, the dissertation offers the concluding remarks.

1.6 List of Papers

The following list gives an comprehensive overview of all scientific papers published by the author that are relevant for this dissertation. Significant parts of this dissertation (across all chapters) were copied from the relevant earable papers listed below and assembled into a coherent monograph structure.

T. Röddiger, D. Wolffram, D. Laubenstein, M. Budde, and M. Beigl. Towards Respiration Rate Monitoring Using an In-Ear Headphone Inertial Measurement Unit. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 48–53. Association for Computing Machinery, 2019

T. Röddiger, M. Beigl, and A. Exler. Design space and usability of earable prototyping. In *Proceedings of the 2020 International Symposium on Wearable Computers*, pages 73–78, 2020

T. Röddiger, M. Beigl, M. Hefenbrock, D. Wolffram, and E. Pescara. Detecting Episodes of Increased Cough Using Kinetic Earables. In *Augmented Humans Conference 2021*, pages 111–115, 2021

T. Röddiger, C. Clarke, D. Wolffram, M. Budde, and M. Beigl. EarRumble: Discreet Hands-and Eyes-Free Input by Voluntary Tensor Tympani Muscle Contraction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021

T. Röddiger, C. Dinse, and M. Beigl. Wearability and Comfort of Earables During Sleep. In *2021 International Symposium on Wearable Computers*, pages 150–152, 2021

S. Hermann, P. Breitling, **T. Röddiger**, and M. Beigl. Cardiopulmonary Resuscitation Support: Comparison of Wrist-, Chest-, and Ear-Worn Devices and Estimation Algorithms. In *2021 International Symposium on Wearable Computers*, pages 28–32, 2021

L. Fang, **T. Röddiger**, F. Schmid, and M. Beigl. EarRecorder: A Multi-Device Earable Data Collection Toolkit. In *Augmented Humans Conference 2021*, pages 286–288, 2021

H. Zhao, **T. Röddiger**, and M. Beigl. AirCase: Earable Charging Case with Air Quality Monitoring and Soundscape Sonification. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 180–184, 2021

T. Röddiger, C. Clarke, P. Breitling, T. Schneegans, H. Zhao, H. Gellersen, and M. Beigl. Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3):1–57, 2022

T. Röddiger, T. King, D. R. Roodt, C. Clarke, and M. Beigl. OpenEarable: Open Hardware Earable Sensing Platform. In *Proceedings of the 1st International Workshop on Earable Computing, EarComp*, volume 22, pages 29–34, 2022

M. T. Knierim, D. Puhl, G. Ivucic, and **T. Röddiger**. OpenBCI + 3D-Printed Headphones = Open ExG Headphones – An Open-Source Research Platform for Biopotential Earable Applications. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–7, 2023

2. Taxonomy of Sensing on the Ears

To establish a foundation for all work presented in this dissertation and advance the understanding of sensing with earables, this chapter initially presents a comprehensive analysis of the different phenomena sensed with earables, based on a systematic literature review of 271 peer-reviewed research articles. Each article reviewed was classified with respect to sensing principles applied or investigated, types of information gained, and purposes to which sensing was used. Through iteration, this process resulted in the development of a taxonomy of *phenomena* sensed with earables. At the lowest level, fundamental phenomena that are directly sensed with sensors placed in or on the ear were identified and characterized. These fundamental phenomena include, for instance, motion, body temperature, and blood perfusion. Other phenomena were identified as indirectly observable and derived from fundamental phenomena, ranging from physiological parameters (e.g., heart rate) and lower-level cues (e.g., earable state; in or out of ear) to conditions (e.g., stress), actions (e.g. gestures), activities (e.g. daily tasks) and other context (e.g. user identity). In total, close to 50 phenomena were identified and categorized. The relationship between higher-level phenomena and fundamental phenomena is demonstrated, and this is related to different sensors and sensing principles that have been investigated for their observation. The result is a taxonomy that is open-ended (new sensors might emerge, and further phenomena explored) but complete in providing a clear end-to-end structure for classification of earable sensing work, from sensors employed to fundamental and higher-level phenomena and their application. This chapter was published as journal paper in "Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies Volume 6 Issue 3" [455]. Also, the related work sections of this dissertation were part of this publication.

2.1 Related Work

A number of recent articles have also reviewed earables research. Plazak and Kersten-Oertel reviewed the properties and affordances of earables and how they are distinct from other wearables, for input and output [379]. Choudhury provided a reflection on earable computing that draws out key opportunities and challenges [102]. The present work, in contrast, is more specifically focused on sensing with earables, and grounded in a systematic literature review. Two other systematic reviews have been published recently. Masè et al. reviewed 39 studies of in-ear monitoring of physiological parameters, focused on temperature, heart rate, and oxygen saturation [298]. Ne et al. also focused on earables for health monitoring but considered a wider range of biosignals, reviewing 92 studies to capture device characteristics versus study outcomes [340]. In contrast, this review presented as part of this dissertation considers earable sensing broadly, not limited to health monitoring, and inclusive also of research that has been less experimental, for example demonstrating novel forms of interaction enabled by earable sensing. The review is organized by phenomena, in four main areas of (i) physiological monitoring and health, (ii) movement and activity, (iii) interaction, and (iv) authentication and identification. For each of the phenomena, a clear definition is provided, and work on how they are sensed and on applications they enable is reviewed. As such, the presented review contributes a uniquely comprehensive survey of the state of the art in earable sensing.

2.2 Methodology

Informed by prior work [248; 51; 163; 499; 75], the systematic literature review was undertaken by collecting and filtering papers from the ACM and IEEE digital libraries using a set of defined inclusion and exclusion criteria and a four-eyes principle, followed by backward chaining with the same criteria applied. This process resulted in 271 relevant articles which were analyzed and clustered based on a newly introduced earable taxonomy.

2.2.1 Paper Retrieval

An initial keyword-based search was conducted on the ACM Digital Library (ACM-DL) and IEEE Xplore (IEEE-X) libraries, which are considered to contain the majority of wearable and HCI publications. The definition of earables was formulated as follows to guide the survey:

> **Earables** are devices that attach in, on, or in the immediate vicinity of the ear to offer functionalities beyond basic audio in- and output.

Research areas that apply to earables but were excluded include voice and audio interfaces as well as the technical design of audio earphones (e.g., speaker and antenna design or noise cancelling algorithms). These topics have been summarised elsewhere (e.g., voice interfaces [428], audio interfaces [143], soundscapes [199], or sonification [250]) or are not specific to the location on the ear (e.g., algorithms to translate speech to text).

The queries listed below were used to match keywords against title, abstract, and author keywords. Keywords were identified by assembling a list that was expanded with additional relevant keywords found in the first 50 papers returned by both libraries. The search resulted in 210 ACM-DL and 695 IEEE-X publications. A final query of both libraries was performed on Jan. 21st, 2022.

```
query target: Title, Keywords, Abstract (ACM-DL) / Document
Title, Index Terms, Abstract (IEEE-X)
keywords: earable(s), hearable(s), ear-worn, ear AND
wearable(s), earbud(s), earphone(s), headphone(s),
earpiece(s), ear(-)mounted, ear(-)attached, ear(-)based
filter: Research Article OR Short Paper (ACM-DL) /
Conferences, Journals (IEEE-X)
```

2.2.2 Selection Criteria, Filtering, and Backward Chaining

As the keyword-based search does not result in a set of papers with clear definition boundaries, explicit inclusion and exclusion criteria were defined to manually select the relevant papers returned by the queries. First, selected papers have to fulfill the basic properties of earables in that:

1. the device attaches in, on, or in the immediate vicinity of the ear; and
2. sensing occurs in, on, or in the immediate vicinity of the ear.

Additionally, articles were excluded that:

1. are not peer-reviewed, e.g. workshop proposals, theses, patents, technical reports
2. are a bigger head-worn or off-body system (e.g., VR headsets)
3. focus on animals
4. are not written in English

The main objective was to show the full breadth and depth of earable research contributions to date. Hence, papers were not excluded based on number of citations, impact factor of the venue, or number of study participants.

The search produced 906 results. Together with a second researcher, the initial results were reviewed separately by reading the titles and abstracts before screening the papers and applying the above criteria. After removing 24 duplicates and one broken cross-site link the articles with positive agreement were selected for the the review (75 ACM-DL, 112 IEEE-X). Performing this step yielded an initial set of 187 papers. Then backward-chaining was applied to the selected papers to account for publications that are not available in the ACM-DL and IEEE-X library or were missed by the keywords, resulting in the inclusion of papers from other publishers including Springer, Frontiers, and MDPI. Together with a second researcher, the references of the papers (4,854 incl. duplicates) were scanned according to the same criteria. All references were split in half and one researcher confirmed the selection of the other. In total, 82 additional papers were identified through this process. Three further papers were added manually that did not appear in the search process, which is a common practice [75]. After going through all papers in depth, 10 additional papers were excluded – five because of severely flawed experiments and five because no evaluation was done but the claims of the paper would demand it. The described procedure resulted in the final set of 271 papers.

Venues with more than three papers were the IEEE International Engineering in Medicine and Biology Conference (EMBC) (N=28), EarComp (N=15), Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT) / Ubicomp (N=14), ACM International Symposium on Wearable Computers (ISWC) (N=13), IEEE Sensors (N=13), IEEE-EMBS BSN International Conference on Wearable and Implantable Body Sensor Networks (N=10), IEEE Transactions on Biomedical Engineering (N=10), MDPI Sensors (N=7), ACM Conference on Human Factors in Computing Systems (CHI) (N=6), ACM Symposium on User Interface Software and Technology (UIST) (N=6), BioCAS (N=5), Mobicom (N=4), and SenSys (N=4).

2.2.3 Analysis and Structuring Approach

To structure the identified papers' content in a unified format, a *Google Sheets* document was assembled. Together with a second researcher, 15 papers were read by each to came up with an initial suggestion for a table structure. This structure was reviewed with four other researchers to distill the final data table (46 columns spanning varying aspects). Then, all papers were split between four researchers to fill the table accordingly.

At the highest level, four main areas of research to group papers were identified which provide a top-level structure. The grouping is pragmatic and based on larger overarching themes. The largest area, in number of articles published, is physiological sensing and health monitoring. Research in this space is pursued across disciplines and has a strong measurement focus but also includes work on detection of distinct phenomena, such as teeth grinding and coughing. Movement and activity forms an area that is more defined by a common ground in activity analysis than any specific application domain, with most of the research stemming from the wearable and ubiquitous computing community. Interaction forms another distinct area, where research tends to be exploratory in pursuit of new means for input and interaction enabled by earable sensing. Authentication and identification is the smallest area but distinct with a research focus on biometrics captured at the ear, including physical properties of the ear itself. Figure 2.1 shows how the four main research areas have evolved over time. The field has grown over the past 20 years with a significant rise in activ-

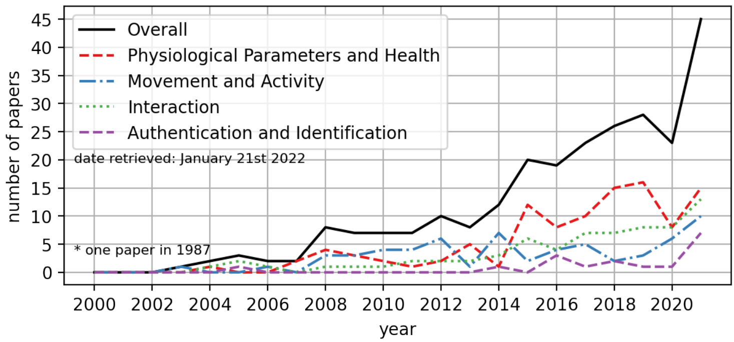


Figure 2.1.: Total number of papers published overall and per research area.

ity in the last 6-7 years. In every of the four main areas identified, most papers were published in 2021, underlining the growing interest in earables.

For the discussion of the state of the art in this dissertation, works with more study participants are prioritised and preliminary findings are signalled. The exact number of study participants of related works are presented in the appendix tables (see section A.1 to section A.15). Works from the same authors with overlapping contents (e.g., follow-up papers) or the same underlying system are not filtered. Instead, overlapping contributions are attributed by citing all relevant papers while highlighting specific contributions through citations of the specific paper.

2.3 Earable Taxonomy

There are many ways in which the research space of earables can be structured, for example by affordances [379] or features of earable platforms [451]. In this dissertation, works are classified by types of sensor and purpose to which sensors were employed. In iteration, phenomena were identified, in the sense of "something that can be observed" as central for structuring the body of work, as it provides the link between sensors as the means for observation, and applications as the target. The use of "phenomena" is comparable to the use of "context" as abstraction in sensing-based applications, however "phenomena" was chosen to better encompass observation of anything from low-level physiological parameter to higher-level condition, event, state, or activity.

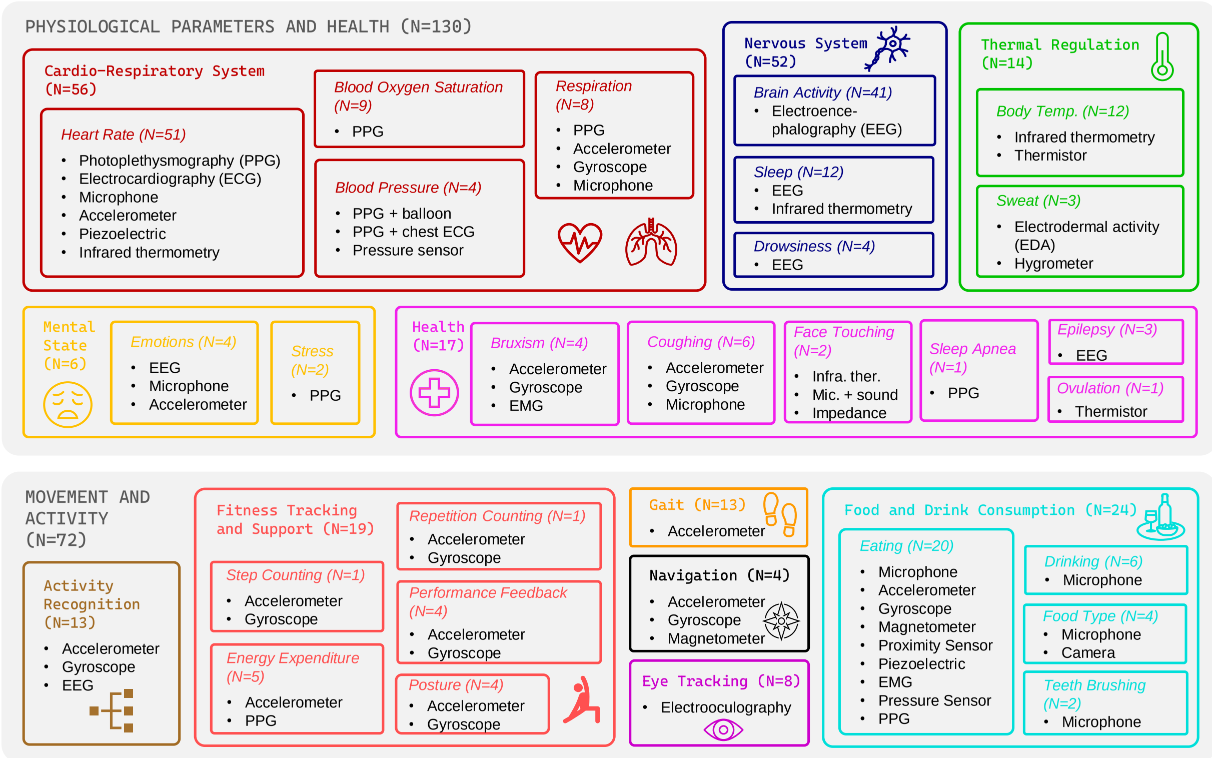
Figure 2.2 provides an overview map of the phenomena identified, clustered into related themes and grouped into four main areas. For example, heart rate, blood oxygen saturation, blood pressure and respiration are all distinct phenomena but clustered as relating to the cardio-respiratory system. For each of the phenomena, the map captures the types of sensors that have been investigated for their observation. The overview map also shows how many articles were found that studied the various phenomena. For example, 49 articles studied earable sensing of heart rate, while other phenomena have only been studied in single works (e.g., detection of sleep apnea, or workout repetition counting).

2.3.1 Definition of Phenomena Sensed with Earables

Table 2.1 provides the list of phenomena distilled from the earable research literature. In many cases, higher-level observations build on lower-level observations that often remain implicit. In developing a taxonomy, all selected articles were analyzed in depth, to clearly identify the actual phenomena observed and phenomena that are derived. For each of the phenomena thus identified, a clear description is provided as point of reference for future research.

In the table phenomena are listed from lower to higher levels of observations, as that allows to show how phenomena build on each other. Phenomena that can be directly captured by a sensor are shown in boldface, and they are referred to as *fundamental phenomena* that enable observation of other phenomena. For example, blood perfusion and cardiac action potential can be sensed directly at the ear, whereas heart rate is derived from lower-level observations.

We also identify categories to capture the principal types of phenomena. A wide range of earable sensing work is, for instance, focused on monitoring of body functions. These are interesting as they focus on physiological parameters as observations at a lower level of abstraction that directly underpin applications concerned with their monitoring. Other categories relate to sensing modalities such as sound, movement and vision, with sound listed at a lower level as it contributes to observation of a wide range of other phenomena. A category of particular note is “Ear” as it encapsulates phenomena that relate to the ear or earable device as such, for example the unique shape of the ear channel, or the manipulation of either the ear or the ear-worn device.



Continued on next page

Continued from previous page

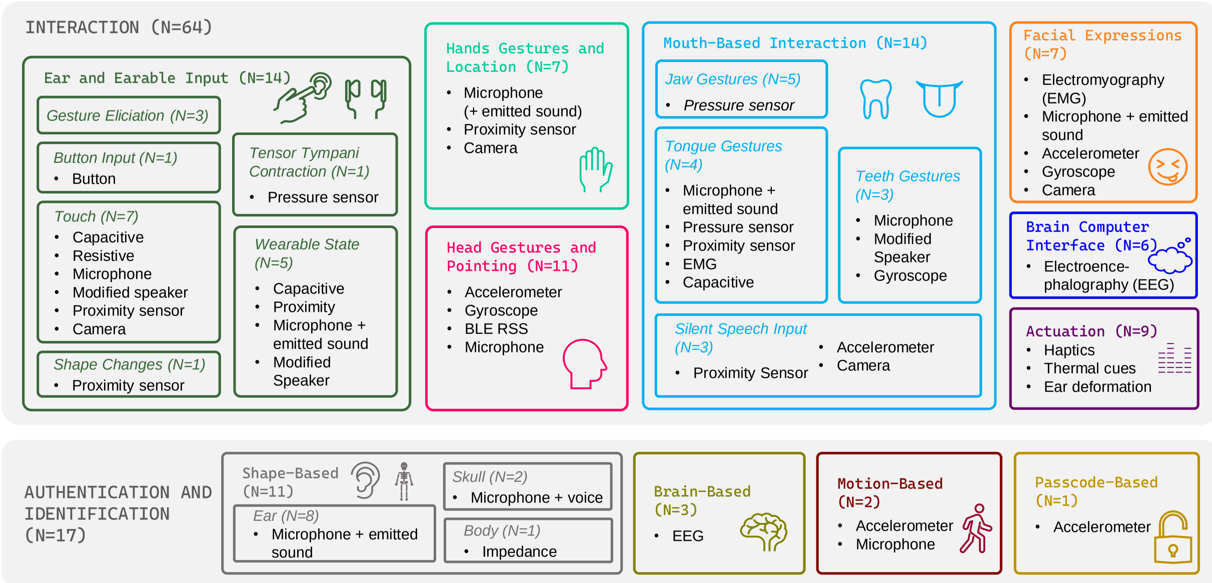


Figure 2.2.: Overview map of earable sensing. The map is organised by phenomena that can be captured with earable sensors. For each phenomenon, the number of articles found in the survey is listed, and the different types of sensor used. Phenomena are clustered in relating to common themes or systems, and grouped into four main categories.

Table 2.1.: Definition of phenomena investigated with earables. Boldface denotes fundamental phenomena that can be observed by sensors and that enable observation of higher-level phenomena.

Category	Phenomena	Description
Body Functions	**Blood Perfusion**	passage of blood through organs and tissue to deliver oxygen and nutrients
	Cardiac Muscle	muscles involved in the contraction of the heart stimulated by cardiac action potential
	Heart Rate	heart contraction frequency varying upon physiological and psychological conditions
	Blood Pressure	pressure created by the heart pushing blood against the walls of the arteries
	Oxygen Saturation	percentage of oxygenated hemoglobin relative to the overall hemoglobin in blood
	Respiration	gas exchange through breathing characterized by inhalation of oxygenated air and exhalation
	Brain Activity	neural activity of the brain emitting electrical signals in response to stimuli or conditions
	Body Temperature	safe range that varies slightly during activities or physiological states (e.g., sports, illness)
	Sweat	water secreted on skin for thermoregulation and in response to psycho-physiological arousal
	Energy Expenditure	energy burned by physical activity and by sustaining human life (e.g., heating the body)
	Sleep, Drowsiness	different stages during sleep, the daily change from wakefulness to sleep, and drowsiness
	Ovulation	can be associated with the highest fertility during the menstrual cycle
Sound	**External Sounds**	sounds occuring externally of the user's body e.g., in public settings by others
	Emitted Sounds	sounds that are emitted by the earable and are then sensed (e.g., sound reflection in ear canal)
	Body Sounds	sounds occuring inside or by the wearer e.g., while chewing or when tapping the ear
Movement / Location	**Motion**	change of position and orientation of an object in space over time, mostly the body
	Navigation	track the movement of a user to compute past and future path or give directions
	Head	movement of the head in different directions limited by the anatomical abilities of the body
	Facial Muscles	contraction of facial muscles to move different parts of the face (e.g., lips, jaw)
	Facial Expressions	conscious and subconscious positioning of facial muscles (e.g., to express emotions)
	Jaw, Teeth, Tongue	conscious and subconscious movement of the jaw, teeth, and tongue (e.g., eating, clenching)
	Eyes	conscious and subconscious movement of the eyes (e.g., gazing, sleep)
	Hands	positioning of the hands and fingers in space over time
Ear	**Shape**	unique shapes of the outer, middle and inner ear and skull around the ear specific to a person
	Deformation	possibility to deform the soft auricle by hand and also ear canal during, e.g., facial activities
	Touch	bring the hands or earable in contact with the skin on and around the ear
	Proximity	distance measured from the ear to other objects (e.g., hands) or ear canal and changes thereof
	Manipulation	manipulation of the ear or earable to perform input
	Earable State	position and status of an earable (e.g., in-/outside the ear, ready for input, etc.)

Continued on next page

Continued from previous page

Mental State	Emotion	feelings and thoughts of a person associated with their physical and psychophysiological state
	Stress	overload of a person's ability to cope with mental or physical demands effectively
Health Conditions	Bruxism	grinding of teeth and clenching of the jaw, many times during the night
	Coughing	ejection of air from lungs with sudden noise to free the lungs from mucus and other particles
	Sleep Apnea	involuntary interruption of breathing during sleep from a few seconds up to minutes
	Epilepsy	commonly associated with episodic abnormal neural activity resulting in, e.g., body shaking
Activities	Posture	static position held by a person while standing, sitting or laying down
	Gait	motion of limbs during locomotion which may be impaired, e.g., due to skeletal malfunctions
	Fitness Activities	activities related to personal fitness or higher physical exertion to track repetitions and execution
	Everyday Activities	activities that occur or are executed daily that can be tracked (e.g., desk work)
	Eating	consumption of foods by bite, chew, and swallow and the detection of foods and intake progress
	Drinking	consumption of liquids (incl. liquid foods) sip and swallow
	Tooth Brushing	brushing of the teeth for oral hygiene to prevent caries
Identity	User Identity	uniqueness of an individual among others based on person-specific properties (e.g., ear shape)
Action	Gestures	movement of parts of the body (e.g., head, hands, jaw) to accomplish input
	Silent Speech	speak words and sentences by moving the mouth and tongue without vocalizing speech sounds
Vision	**Body Appearance**	looks of the user's body which give information about presence, state, or identity
	Object Appearance	looks of surroundings which give information about presence or state of objects

In the ubiquitous computing field, lower-level observations are often referred to as “cues” that contribute to inference of “higher-level context”. Similar relationships are visible in the presented schema, where other categories such as mental states, health conditions, activities and identity describe higher-level contexts. However, across the examined body of work, phenomena of interest are viewed at different levels of abstraction. Therefore, the dissertation avoids any layering into cue versus context. The same phenomenon may be considered low level in one application and high level in another.

Figure 2.3 shows the relationships identified between phenomena. Fundamental phenomena that can be directly sensed are shown as grey boxes, with other phenomena in blue. For a range of phenomena, the relationship appears straightforward, for example with posture or gait derived from motion. However, phenomena build also in less obvious ways on each other. Observation of heart rate, for instance, can be based on observation of cardiac muscle activity, blood perfusion and sound but in turn also contributes to observation of a spread of phenomena including blood pressure, respiration and fitness activity. Another example is ear canal shape deformation which feeds into detecting jaw, teeth, and tongue movements which enables the observation of higher-level phenomena such as eating, bruxism or silent speech detection. Simultaneously, the ear’s distinct shape can also reveal the user’s identity. While changes of ear canal shape can be quantified directly using piezoelectric or pressure sensors, it may also be sensed indirectly by measuring motion changes over time or by characteristic sound reflections emitted in the enclosed ear canal.

2.3.2 Fundamental Phenomena

Figure 2.3 highlights how the observation of a wide range of phenomena is grounded in a relatively small set of fundamental phenomena. These fundamental phenomena are sensed either directly or indirectly from different sensors attached in, on or in the immediate vicinity of the ear. However, the phenomena labelled as fundamental may also be inferred from other fundamental phenomena in addition to being (in)directly sensed. For instance, “Touch” is classified as a fundamental phenomenon because it can be captured directly with specific sensors (see Table 2.3). However, "Touch" can also be inferred indirectly, for example from sound captured with a microphone or visual appearance of body and hand captured with an ear-mounted camera.

Table 2.2 provides a detailed definition for the identified fundamental phenomena, explaining the physiological mechanisms on which their observation is based. The description makes reference to the anatomy of ear and head, for readers should refer to Figure 2.4. The table provides a comprehensive overview of the foundations on which earable sensing is based. Note also specifics, for instance how external, emitted and body sound leverages earable sensing differently.

2.3.3 Sensing Principles

Table 2.3 provides a list of the different types of sensors that were reported in the earables literature, grouped into general categories. For each sensor, the dissertation provides a description of the sensing principle for reference. Each sensor directly relates to sensing at least one of the fundamental phenomena, however some fundamental phenomena can be sensed with a wide variety of sensors. For example, ear canal deformation can be sensed using a barometer to detect the in-ear pressure, through proximity sensors which measure the in-ear distances, or through accelerometers and gyroscopes as the ear canal deforms during jaw movements.

The list of sensors also completes the circle to the overview map of the taxonomy (Figure 2.2), which mapped out the range of phenomena with the different types of sensors used for their observation. The taxonomy provides a clear structure based on sensors, fundamental phenomena, higher-level phenomena, and their relationships. The remainder of this dissertation builds upon the presented taxonomy. For future research, the taxonomy should provide a reference scheme for new work in the field, for which the taxonomy will naturally be extensible for other sensors and phenomena of interest.

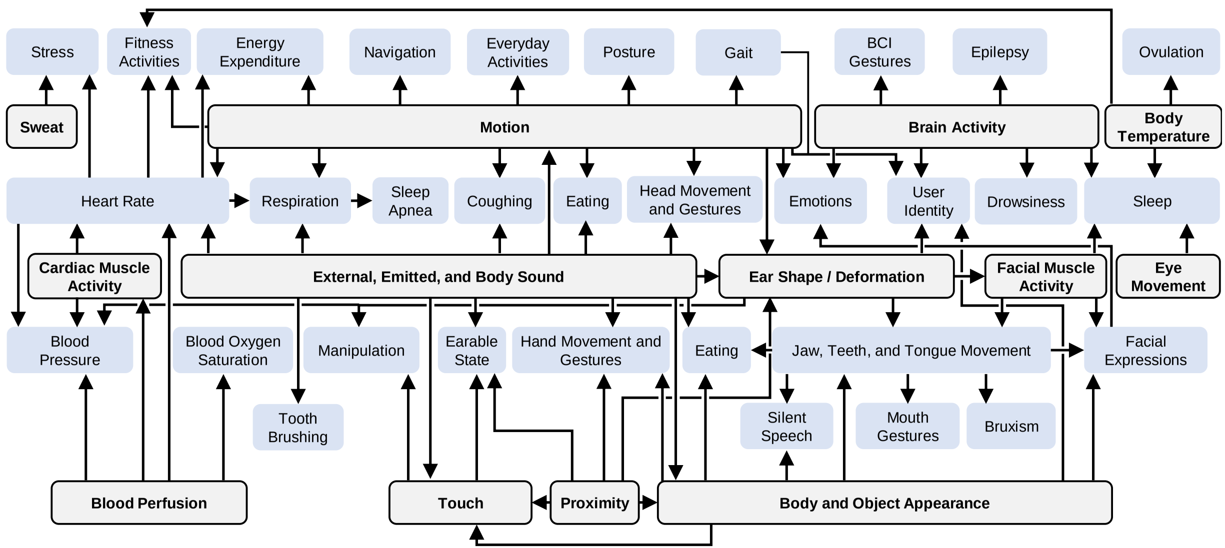


Figure 2.3.: Flow diagram showing how different phenomena can be inferred. The arrows indicate how lower-level phenomena support higher-level observation. The grey boxes represent "fundamental" phenomena that can be directly sensed and from which all other phenomena can be derived.

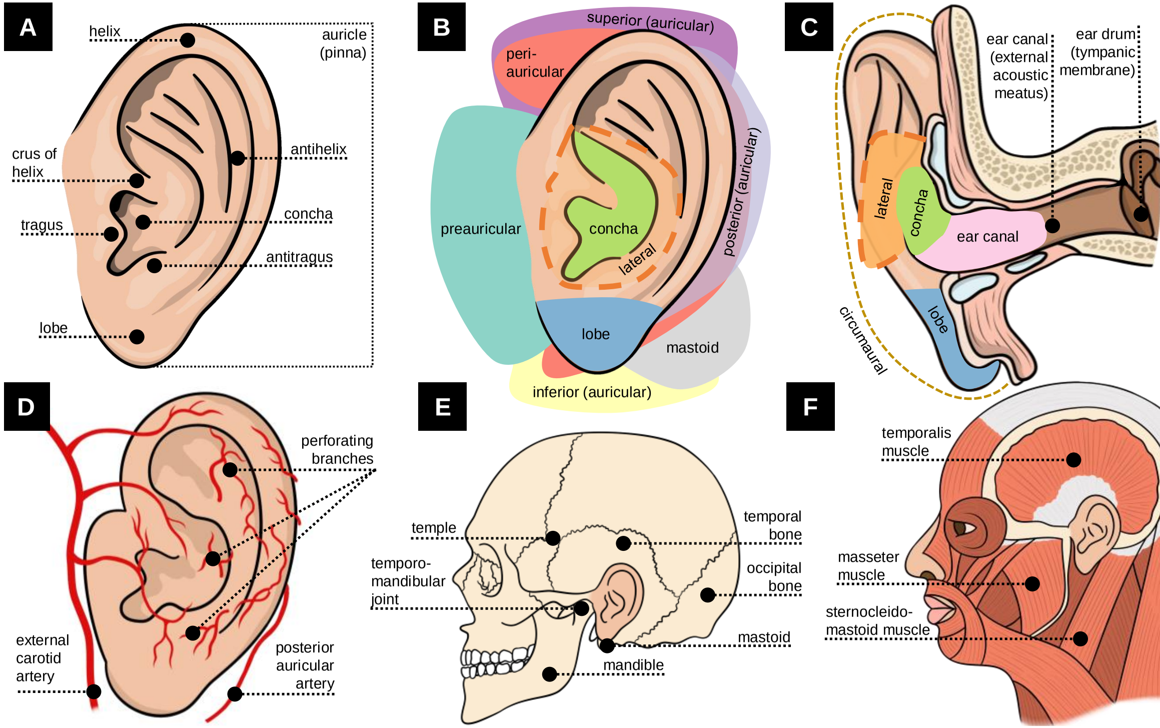


Figure 2.4.: (A - C) ear anatomy and unifying positioning terminology; (D) arterial network; (E) skull bones; (F) muscles.

Table 2.2.: Fundamental phenomena from which all other phenomena can be derived on the earable platform.

Fundamental Phenomena	**Description / Underlying Mechanisms**
Blood Perfusion	The ears are characterized by thin tissue and visible blood vessels (see Figure 2.4 D) enabling the observation of cyclic blood flow and related properties [53]. The perfusion of the middle ear is excellent in comparison to the peripheral perfusion of other body parts [469].
Cardiac Muscle Activity	Due to the conductive characteristics of the body, the cardiac action potential created by the heart is propagated throughout the body up to the ear [502; 173; 419].
Brain Activity	The area around the ear, the concha, and the ear canal are closely located to the brain, which allows capturing its electric activity, commonly resulting in sinusoidal waves that are also called brain waves [446; 224; 57]. It provides access to the brain's response upon auditory and visual stimulation [224; 342]
Body Temperature	The close proximity to the carotid artery results in the tympanic membrane having a temperature close to core body temperature [62] (see Figure 2.4 C and D). Additionally, the ear canal and an earbud create a confined space in which temperature stabilizes [284].
Sweat	The area around the ear has high sweat-gland density relating to stress and physical exertion [426]. Sweat gland activity is not symmetric and different between both ears [377].
External / Emitted / Body Sound	Activities occurring close to the ear (e.g., chewing sounds or tapping around the ear) are transmitted by body sounds, or bone conduction [13; 494]. The cavity created by the ear and an earable generates a natural echo chamber that amplifies body-internal sounds [257; 359], while external sounds are dampened [364; 359]. Sounds that are actively emitted from a device at the ear result in characteristic sound reflections that are utilised by different phenomena, including detection of ear canal deformation [11]. Compared to smartwatches and smartphones, earables are less susceptible to motion-induced sound artifacts [318].
Motion	The ear provides a robust and stable attachment point [84] with few vibrations and random movement artifacts [136] when detecting motions across the body and at the ear. This includes motion induced by the ear canal deforming (e.g., during facial expressions [472]).

Continued on next page

Continued from previous page

Facial Muscle Activity	Sensing of facial and neck muscles can be achieved via electrical potential changes in the area around the ear [304], which is closely located to the temporalis, masseter, and steroncleodmastoid muscle (see Figure 2.4). Facial muscles deform the ear canal [16].
Ear Shape / Deformation	The fine structures of the ears are unique enough between different users (see Figure 2.4 A) for the purpose of biometrics [378; 37]. Additionally, the ear canal deforms during head, face, mouth, teeth and jaw motions and muscular activity [16; 47; 162; 354], e.g. upon movement of the temporomandibular joint (see Figure 2.4 E).
Touch	The unique structure of the ear and the earable itself create a surface for interaction and are easily reached by the hands which affords touches by the user [272; 241].
Proximity	The ear offers a fixed reference point from which distance to external objects can be measured, or their presence inferred [312]. In addition, in-ear based proximity sensors can be used to detect ear canal deformation [47].
Eye Movement	The standing potential of the eyes and, upon movement of the eyes, changes thereof can propagate to the ear [57].
Visual / Object Appearance	The location at the ears can capture the field of view of the wearer and also the broader area around them which contains visual information about the appearance of the surroundings [273] and can also determine touch [241].

Table 2.3.: Sensors can quantify or measure different fundamental phenomena that were identified in Table 2.2 and Figure 2.3. Some fundamental phenomena can still be inferred from other fundamental phenomena.

	Sensor	Description	Fundamental	Inferred Fund.
Motion	Accelerometer	Inertial acceleration of a body on one or more axes to measure body motion changes and ear canal deformation motions relating to facial muscle activity.	Motion	Ear Deformation, Facial Muscle Activity
	Gyroscope	Angular rotation of a body on one or more axes to measure body motion changes and ear canal deformation motions relating to facial muscle activity.	Motion	Ear Deformation, Facial Muscle Activity
	Magnetometer	Magnetic field on one or more axis relating to body motion changes	Motion	-
Audio	Microphone	Sound transmitted through air or occurring within the body of the user. Such sounds can relate to touching.	External and Body Sound	Touch, Motion
	Microphone + emitted sound	Reflected sound signals of an emitted sound (e.g., constant tone) corresponding to ear shape and deformation with facial muscle activity	Emitted sound	Ear Shape / Deform. Appearance, Facial Muscle Activity
	Modified Speaker	Modification of a speaker to sense sounds that actuate the speaker's membrane. Such sounds can be leveraged to sense touch.	External and Body Sound	Touch
Optical	Photoplethysmography (PPG)	Absorption of emitted light corresponding to blood perfusion (possibly at different light frequencies).	Blood Perfusion	Cardiac Muscle Activity
	Infrared Thermometry	Temperature based on thermal radiation without physical contact, commonly obtained at the tympanic membrane.	Body Temperature	Blood Perfusion, Appearance, Touch
	Proximity Sensor	Distance to objects in close proximity with no physical contact to observe appearance which can sense deformation and touch.	Proximity	Ear Deformation, Touch, Appearance
	Camera	Capture images to sense visual appearance of surroundings (e.g., hands), contact to the skin, or deformation of the face.	Visual Appearance	Touch, Facial Muscle Activity
Biopotential	Electroencephalography (EEG)	Electrical potential changes of the brain created by neurons in response to external stimuli (e.g., visual) or internal processes (e.g., sleep, emotions).	Brain Activity	Motion
	Electrooculography (EOG)	Movement of the eye's standing potential (dipole) and eye lid by change of potential around the eyes measured at the ear.	Eye Movement	-
	Electrocardiography (ECG)	Electrical potential of the heart produced by cardiac muscle contraction that propagates through the body.	Cardiac Muscle	-
	Electromyography (EMG)	Electrical potential generated by muscle cells during contraction to sense the activity of muscles around the ear and of the face.	Facial Muscles	-
	Electrodermal Activity (EDA)	Changes of skin lead resistance upon secreted sweat in response to psychological or physiological arousal.	Sweat	-
	Impedance	measures the impedance of current influenced by tissue	Touch	Appearance

Continued on next page

Continued from previous page

Environmental	Thermistor	Temperature as resistance change, commonly attached on the object to be sensed, or inside a confined space such as the ear canal.	Body Temperature	-
	Hygrometer	Measures humidity as the concentration of vaporized water in air which corresponds to sweat secreted by the skin in the ear canal	Sweat	-
	Barometer	Measure air pressure in the ear canal to sense deformations that corresponds to facial muscle activity.	Ear Deformation	Facial Muscle Activity
Electrical	Piezoelectric/ -resistive	Sense change of electric charge / resistance upon mechanical stress relating to motion of ear deformations and facial muscle activity.	Touch, Motion, Ear Defor.	Facial Activity
	Capacitive	Sense change by capacitive coupling to a conductor or materials with different dieletric properties, such as the finger.	Proximity, Touch	-
	Button	Outputs a binary state or pressure force level.	Touch	-
	BLE RSS	Bluetooth signal strength between two devices.	Proximity	-

3. Physiological Parameters and Health

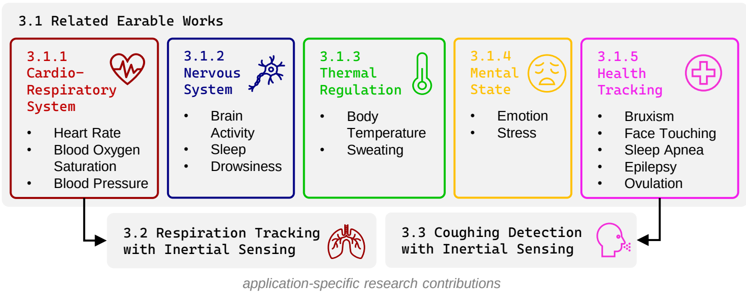


Figure 3.1.: Structure of the “Physiological Parameters and Health” section according to different functions of the human body. The sections "Respiration" and "Coughing" are separated out because of the sensing contributions and studies conducted in this dissertation.

This chapter focuses on physiological parameters and health applications related to ear-based sensing. Figure 3.1 gives an overview of the structure of the chapter, which includes related works and application specific contributions.

Section 3.1 summarizes related earables works according to various functions of the human body. They describe the technical realization and applications of various bodily phenomena detected by sensors attached to the ear motivated by tracking and maintaining personal health.

Two body functions are separated out into their own sections because of the application-specific contributions of this dissertation. Section 3.2 introduces respiration rate sensing based on inertial data and also related works. In section 3.3, episodes of increased cough detection is presented based on inertial sensing and also related works are summarized.

3.1 Related Earable Works

The following sections are structured according to functions of the human body, as illustrated Figure 3.1. Respiration and coughing are separated out into their own sections 3.2 and 3.3, respectivley.

3.1.1 Cardio-Respiratory System

The cardio-respiratory system resembles a close coupling process between two biological phenomena: blood flow and breathing cycles. The following sections will introduce the technical realization and applications of heart rate, blood oxygen saturation, and blood pressure sensing on the ear. Respiration-related sensing is introduced section 3.2 in line with the approach for respiration rate tracking introduced in this dissertation.

3.1.1.1 Heart Rate. Heart rate (HR) is an indicator of the cardiovascular and the autonomic nervous system and, therefore, a vital sign that is influenced by physical fitness, diets, and the overall health [464]. It describes the frequency at which the heart contracts and relaxes. A typical heartbeat consists of multiple characteristic waves (most importantly P-, QRS-, and T-waves [173]). Typically, the heart rate is identified from the R-wave and reported in beats per minute (bpm), with adults having an average resting heart rate of 60-100 bpm [501]. Heart rate variability (HRV) is the variability in the beat-to-beat time intervals which can predict cardiovascular diseases and mortality [458; 457]. HRV is reported in milliseconds (ms), with adults typically having an average resting HRV of approximately 20-200 ms [349; 251].

Heart Rate - Sensing. Table 3.1 compares seven earable heart rate sensing principles based on the results of 44 studies (for details, see appendix A.1). The different sensor locations, experimental protocols, experimental conditions, and performance metrics reported across papers only allow an ordinal comparison of accuracy and robustness (low, medium, high). In this dissertation, “high accuracy” is defined as medical-grade accuracy in the resting condition (e.g., mean error < 10% [24; 353]). Robustness refers to the stability against motion artifacts. The following paragraphs will describe the advantages and disadvantages of the different sensing principles in further detail.

Table 3.1.: Comparison of heart rate sensing principles. Accuracy based on comparison to medical gold standard under resting conditions. Robustness refers to the robustness against body movements. Accur. = Accuracy, Robustn. = Robustness, PPG = Photoplethysmography, ECG = Electrocardiography, Mic. = microphone, Acc. = Accelerometer, N = Number of studies

Sensor	**Accur.**	**Robustn.**	**Advantages**	**Disavantages**	**Best Loc.**	**N**
PPG	high	moderate	supports pulse oximetry, blood pressure, and respiration rate	sensitive to motion but can be filtered up to some degree	ear canal	29
ECG	high	moderate	gold standard for HR HRV, most detailed heart activity	does not work for everyone, obtrusive, requires multiple electrodes	ear canal, mastoid	8
Mic.	medium	low	off-the-shelf	sensitive to motion but can be filtered up to some degree	ear canal, circumaural	5
Acc.	medium	low	off-the-shelf, filter movement supports activity tracking	highly sensitive against movement	posterior	2
Piezo	high	low	configurable ear shapes, robust against temperature	highly sensitive against movement, requires pressure to the skin	ear canal	1

Photoplethysmography (PPG) measures the blood volume change by illuminating the skin and then tracks changes in the reflected or transmitted light. The proximity between the brain and ears offers an arterial network that is ideal for heart rate sensing in comparison to other locations that are subject to peripheral perfusion (see Figure 2.2). At the same time, PPG affords sensing other phenomena such as blood pressure (see subsubsection 3.1.1.3), blood oxygen saturation (see subsubsection 3.1.1.2), and even respiration (see section 3.2). In general, a PPG sensor location in the ear canal is preferable even though it may be sensitive to jaw motions [475; 137]. Other non-ear-obstructing locations with sufficient accuracy are at the tragus, the ear lobe, and the area posterior to the ear (see appendix A.1). Accuracy-wise, earable PPG devices can meet official medical standards in resting and moving conditions. Most resting state studies reported error scores of less than 1 bpm mean error [422; 459; 137; 372; 469; 470; 476; 475; 474] which even outperforms wrist-worn PPG [392; 376]. Within-subject studies showed that PPG performance decreases with motion artifacts introduced by body movement [498; 137; 422; 382; 459], facial movement (e.g. talking) [286; 184; 382], or music listening [381]. While several studies report accuracy scores of less than 10% mean error under motion noise [372; 286; 422; 422], other studies exceed the acceptable range of medical standards (e.g., [137]). Two promising pathways for reducing earable PPG motion artifacts are the use of accelerometers [372; 283; 90] and machine learning based calibration procedures [439; 283; 502].

Electrocardiography (ECG) measures the contraction of the cardiac muscle and the resulting electrical activity with electrodes on the skin surface [431]. ECG provides the highest resolution of heart rate activity and is, therefore, considered the gold standard for conventional medical measurement. A handful of studies on earable ECG reported acceptable performance (see appendix A.1) with peak delays of only 50 ms [438]. The ear canal was recommended as the best location for earable ECG [477]. Still, some evidence exists that ECG waves can also be measured at the mastoid [209; 502] and posterior ear position [89]. However, Jacob et al. [209] could not identify the fundamental heart beat frequencies for 6 out of 13 study participants. Additionally, multiple ECG-electrodes are required on the skin (not necessarily on the head), which leads to the conclusion that earable PPG offers significant advantages over ECG with regards to obtrusiveness, generalizability, and possible accuracy.

Microphones [296; 128] record air-conducted (and by modification also bone-conducted [160]) sound pressure waves elicited by mechanical pulsation of the ear-canal blood vessels. The recorded audio signal is processed with filtering algorithms for low (<24Hz) and recurring frequencies [128; 345], or other denoising algorithms [70; 296]. One advantage is that microphones are commonly built into commercial earbuds. Fan et al. [128] developed a device that plugs in between a smartphone and any off-the-shelf headphone to derive heartbeats from small voltage changes. Most studies report errors below the 10% range (appendix A.1). However, the performance of microphone-based heart rate sensing is generally lower than earable PPG. Moreover, it has low robustness against motion artifacts (ME=7.5 bpm) [345]. As a result, microphones should only be considered as a cost-efficient and off-the-shelf alternative to PPG sensors.

Accelerometers measure heartbeats through recurring mechanical vibrations that result from the blood volume change [184; 355]. Similar to microphones, they are built into several off-the-shelf earables (e.g., eSense [236]). He et al. [184] found high regression coefficients for the R- and J-waves but lower coefficients for the PPG-measured stroke volume. Such accelerometers are also useful for filtering out motion artifacts of other heart rate sensors [372; 283; 90; 183] (see above).

Moreover, Park et al. [362] used piezoelectric sensors for measuring the heart rate via the variance of surface pressure in the ear canal. The authors found showed high accuracy scores on a large sample size when the user is at rest.

None of the sensing principles have high robustness against body movements. PPG and ECG are evaluated as moderately robust because they had less decline in accuracy compared to other sensing principles and still enabled an overall identification of the heartbeats. Overall, a combination of PPG sensors in the ear canal and accelerometers that control for motion artifacts seems most promising for earable heart rate sensing.

Heart Rate - Applications. Most earable heart rate publications were motivated by the possibility to continuously monitor cardiovascular functions (e.g., [381; 362; 173]). Other use cases are monitoring of stress [283; 439; 439; 169], see subsubsection 3.1.4.2, energy expenditure [258], see subsubsection 4.1.2.2), and exercising [460; 345]. Heart rate measured at the ear also gives insights into respiration-related events (see section 3.2 and subsubsection 3.1.5.3).

3.1.1.2 Blood Oxygen Saturation. Delivering oxygen-bound hemoglobin to different cells of the body is vital to human life [485]. The proportion of oxygenated hemoglobin (saturated) to the total amount of hemoglobin molecules in the arteries (saturated and unsaturated) is defined as blood oxygen saturation, and is reported in percent (%).

Non-invasive methods commonly measure the peripheral oxygen saturation (SpO_2) based on computing the difference of absorbed light at two wavelengths by Photoplethysmography (PPG). In medicine, the earlobe is a popular and reliable location to obtain transmissive PPG-based SpO_2 measurements as light can be emitted on one side of the ear, whereas the absorption is determined on the other side of the tissue [485]. However, transmissive PPG is limited to the auricle, as the emitting LED and the light sensor are placed on opposing sides of the skin. Therefore, earables commonly use reflective PPG sensors that measure the amount of light reflected back to the emitting probe.

Appendix A.2 shows that acceptable accuracy can be achieved by reflective PPG at rest on the earlobe [286], and in the ear canal [53; 137] based on ground truth finger SpO_2 measurements ($< 2\%$ error according to FDA [141]). Besides, PPG sensors posterior of the ear could also measure reliable values in a range between 70% - 100% evaluated according to gold standard arterial blood gas values (SaO_2) [71]. PPG to sense blood oxygen saturation suffers from the same problems as described in the previous heart rate section (see subsubsection 3.1.1.1), such as degraded performance by motion artifacts when walking or during jaw movement [137; 475].

Based on the blood oxygen saturation of the wearer, it was suggested to perform sleep apnea detection [469], see subsubsection 3.1.5.3. Other possible applications include vital signs tracking and alerting [40; 286; 53] foreseeing diseases [475; 168], and unobtrusively monitoring the oxygen dosage of patients [249].

3.1.1.3 Blood Pressure. The circulatory system transports nutrients to all parts of the body while the pressure generated by the heart pumping plays a decisive role as the driving force. This pressure is known as blood pressure [309]. High blood pressure (hypertension) is a known risk factor of cardiovascular diseases and even death which makes it an important vital sign to track [352]. The blood pressure during systole (contraction) and diastole (relaxation) is called systolic blood pressure (SBP) and diastolic blood pressure (DBP), respectively. Historically, medicine reports blood pressure in millimeters of mercury (1 mmHg = additional pressure by 1 millimeter of mercury).

In related work, Teng and Zhang [445] showed a strong relationship between blood pressure and the time of pulse propagation from the heart to other locations of the body. Therefore, a PPG sensor worn on the ear and a traditional chest-worn ECG makes it possible to compute the Pulse Transit Time (PTT) and hence, blood pressure [168]. Selvaraj [417] demonstrated by regression analysis that PPG measured on the earlobe gives about similar results as the finger to compute PTT yielding a weak correlation between PTT and SBP / DBP ($r < 0.3$, see appendix A.3).

In a concept study it was suggested to measure blood pressure in the sealed ear canal based on pressure sensing with initially encouraging results [500]. The feasibility and precise performance of the principle remains to be evaluated.

The aforementioned approaches prerequisite chest-worn ECG or sealing the ear canal which reduces everyday wearability. Bui et al. [76] introduced *eBP* to measure blood pressure using an ear-worn device only. Very similar to cuff-worn blood pressure sensing [154], eBP places an inflatable balloon in the ear canal of the user. An evaluation with 35 users showed that eBP yields an average error of 1.8 mmHg and -3.1 mmHg and a standard deviation error of 7.2 mmHg and 7.9 mmHg for SBP and DBP, respectively). Though these results are promising, the acceptable accuracy for at-home blood pressure devices was defined to be ± 3 mmHg in other works [194].

The primary purpose of ear-worn blood pressure is a frequent assessment throughout the day as diseases may be episodic and related to specific activities [76; 168; 417]. Compared to traditional cuff-based monitors, ear-worn blood pressure monitoring minimizes the impact during regular activities while maximizing comfort [76; 417].

3.1.2 Nervous System

The nervous system sends and receives electrical and chemical signals to control body functions and cognitive processes [333]. The resulting electrical fields can be measured on the skin surface by Electroencephalography (EEG) to extract a user's response to external stimuli internal states. Such states include sleep stages and the sleep-wake cycle. The following two sections will introduce earable brain activity sensing and sleep tracking.

3.1.2.1 Brain Activity. Sensing the brain activity of a person based on standard testing protocols is used for diagnosing a number of neurological and psychological disorders [334; 175]. Additionally, brain activity sensing enables higher-level applications such as sleep (see subsubsection 3.1.2.2) and emotion (see subsubsection 3.1.4.1) tracking, seizure detection (see subsubsection 3.1.5.4), brain-based authentication and identification (see section 7.4), as well as brain-computer interfaces (see subsection 5.1.6).

Brain Activity - Sensing. EEG measures the electric field potential caused by characteristic brain rhythms (also called brain waves) that correspond to a user's internal state or response to external stimuli [60]. The standard procedure to capture brain waves is Electroencephalography (EEG). Conventional EEG relies on a full-scalp setup worn all around the user's head which is obtrusive and not easily portable. In contrast, a smaller number of electrodes in (e.g., [160]) or around (e.g., [57]) the ear can still capture a subset of interesting brain waves. Generally speaking, related ear EEG works have shown that in comparison to scalp EEG, brain activities primarily occurring in the temporal and occipital lobe around the ear can be recorded successfully [240; 57; 58]. Overall, the reduced size of earable EEG compared to conventional EEG results in several advantages by being more discreet, unobtrusive, robust, user friendly, and feasible [278; 57].

Slight performance differences with regards to impedance and usability aspects specific to the possible electrode positions and accompanying form factors have emerged. For example, 9 out of 10 devices in the posterior and periauricular region have a generic form factor that fits on the skin around the ear.

Similarly, 10 out of 12 devices in the ear canal are implemented as generic soft earplugs. In contrast, 19 out of 21 concha-placed ear EEG devices are custom fit to the wearer as the unique structure of the concha creates a challenging fit across users. Still, electrodes placed in the concha were more prone to lose contact than in the ear canal [228]. Furthermore, EEG placed inside the ear leverages it as a mechanical anchoring point (e.g., [160; 314; 58]), whereas EEG around the ear often demands gluing the electrodes onto the user's skin (e.g., [57; 348; 295]). While generic in-ear EEG can fit many different users, there is some preliminary evidence that custom-fit earplugs are more sensitive than generic ones [239; 233]. At the same time, generic ear EEG worn around the ear was found to be more sensitive than generic in-ear EEG [52].

Across ear EEG locations and styles, 23 dry and 21 wet electrode setups were identified. Dry electrodes improve the comfort of the user and simplify attachment. However, they have reduced impedance or sensitivity depending on the electrode material and compared to wet electrodes [225; 34; 223]. Additionally, the presence of cerumen was found to increase dry-contact impedance by 86% [370]. In general, the close vicinity of the ear to facial muscle potentials [45], as well as eye movement and blinking artifacts [234; 125; 57] can result in noise which was recommended to filter out in related earables works.

As to be expected, no works looked into the feasibility of ear EEG on the auricle. Overall, the different design aspects of ear EEG allow to make trade-offs between form factor, fit, ease of application, comfort, and desired accuracy.

Brain Activity - Applications. Table 3.2 gives an overview of earable EEG implementations and lists standard protocols that were conducted to show the general feasibility of EEG on the ear (see appendix A.6 for details). It also links to other sections that summarize more concrete applications. Generic, dry ear EEG is most generally applicable in day-to-day usage, while custom-fit ear EEG devices with wet electrodes are more likely to be relevant for clinical usage.

In EEG research and clinical practice, auditory, visual, and somatosensory (i.e., haptic) cues are applied to trigger an expected response of the patient's brain [410]. As it stands, little insights are available about the response to

Table 3.2.: Standard EEG protocols that have been evaluated on the ear using generic form factors that can fit any ear and devices that have to be custom fit to the user. Ear EEG applies wet and also dry electrodes. The feasibility of standard EEG protocols shows the general feasibility and enables higher-level applications. In brackets is the number of studies that confirmed the paradigm. A detailed overview of ear EEG papers and placements can be found in appendix A.6.

Electrodes		**Auditory Stimulus**				**Visual Stimulus**		**Task**
		P300	MMN	ASA	ASSR	P300	SSVEP	AAR
Generic	Dry	✓(3)	✓(1)	-	✓(7)	✓(1)	✓(4)	✓(5)
	Wet	✓(3)	-	✓(1)	✓(4)	✓(3)	✓(6)	✓(4)
Custom	Dry	-	✓(1)	✓(1)	✓(6)	✓(1)	✓(5)	✓(6)
	Wet	✓(3)	✓(2)	-	✓(6)	✓(2)	✓(3)	-
Example Application		psychological / neurological disorders [175]	psychological / neurological disorders [334]	spatial audio hearing aids [139]	brain-computer interfaces [190]	psychological / neurological disorders [175]	brain-computer interfaces [5]	sleepiness / drowsiness detection [57]

Note: Other applications based on EEG are introduced in subsubsection 3.1.2.2 (sleep), subsubsection 3.1.4.1 (emotions), subsubsection 3.1.5.4 (seizures), section 7.4 (brain-based authentication and identification), and subsection 5.1.6 (brain-computer interfaces). P300=Positive deflection in brain potential; MMN=Mismatch Negatively; ASA=Auditory spatial attention; ASSR=Auditory steady state response

somatosensory stimuli. However, it was initially shown that skin impedance decreases in response to tactile stimulation [370]. Nonetheless, a variety of auditory and visual stimulus paradigms have been confirmed to be feasible in and around the ear, which will be described in the following.

In a clinical diagnostic context, the P300 response has been associated with dementia, schizophrenia, anxiety disorders, and more [175]. Presenting a deviant stimulus among continuous auditory or visual signals triggers a so-called transient P300 response which is a positive deflection in brain potential approximately 300ms after the deviant stimulus is presented [437]. Auditory P300 was confirmed by multiple earable studies even several hours after initial attachment [57]. Visual P300 with ear EEG was visible in response to letters [58], words [348], symbols [278], LED lights [240; 160; 159], and black/white checkboards [166]. Similarly, Mismatch Negativity (MMN) has clinical relevance for diagnosing schizophrenia or aging [334] and represents a negative EEG amplitude deflection. It targets the lower-level discrimination abilities without the user having to actively focus their attention on a stimulus. Kappel et al. [228] found that a reference electrode on one ear is required to capture a significant MMN response on the other. As such, visual attention or resting may be predicted directly from ear EEG at >70% accuracy [371]. In sum, earables potentially give easy access to diagnosing or monitoring neurological and psychological disorders without the need for a full-scale scalp EEG setup.

Presenting auditory and visual stimuli at a high repetition rate creates an overlap of transient responses, the so-called steady-state response. As the steady-state response depends on the frequency of the presented stimulus, it is ideal for selection tasks of brain-computer interfaces (see subsection 5.1.6). The auditory steady-state response (ASSR) was confirmed with ear EEG in and around the ear with amplitude modulation frequencies ranging from 40 Hz [278] up to 90 Hz [224]. The steady-state visual evoked potential (SSVEP) was also confirmed in and around the ear at many different frequencies ranging from 5.4 Hz [266], up to 20 Hz [240].

Extracting the direction of a user's auditory attention creates a compelling use case for hearing aids that could amplify the sounds from an attended source [139]. In that regard, the auditory spatial attention (ASA) tests the response of a person attending to simultaneous sounds coming from different directions [101]. Wet periauricular ear EEG could achieve similar performance to scalp EEG [59]. The feasibility of ASA in the ear canal with dry electrodes depended on the positioning of the reference electrode [139].

The alpha attenuation response (AAR) describes an increase of alpha frequency power in the EEG signal once a person closes their eyes. The alpha power (brain waves at 8 - 12 Hz) can be associated with the sleepiness of a person [430]. It is well possible to observe a statistically significant alpha attenuation response with in-ear and around-the-ear electrodes. Alpha power increase can also be associated with drowsiness [57], see subsubsection 3.1.2.3.

3.1.2.2 Sleep. It was reported that up to 40% of the U.S. adult population struggles with sleep annually which results in morbidity and mortality [201], which makes sleep an interesting parameter to track. Sleep can be divided into four reoccurring stages (N1, N2, N3, REM), which are repeated up to six times per night [369]. Each stage is characterized by physiological patterns across the body, that are commonly analyzed in professional sleep labs by polysomnography (includes scalp EEG, EMG, nasal airflow, etc.). State-of-the-art sleep stage classification from full-scale polysomnography achieves up to 97% F1 score on diverse datasets [167]. In contrast, earable computing research aims to reduce the number of required sensors to perform sleep analysis from biopotential signals at the ear, making it feasible even at the patient's home [337].

Automatic sleep stage classification from biopotential signals at the ear ties in deeply with the general feasibility of sensing brain activity (see subsubsection 3.1.2.1), facial muscles (see subsection 5.1.5), and movement of the eyes (see section 4.2). Sleep stage prediction in the ear canal (see appendix A.5) was performed with wet electrodes using different evaluation strategies and classifiers. Accuracies between 66% and 95% in comparison to gold standard polysomnography were achieved [337; 342; 315]. Sleep latency (the time it takes to fall asleep) was predicted at less than three minutes error [10]. Based on the presence of sleep spindles in ear-EEG, it was initially shown, that sleep staging with dry electrodes may be feasible [316; 293]. With data captured over 80 nights from 20 participants, Mikkelsen et al. [317] then showed that custom-fit, dry ear-EEG allowed automatic sleep scoring at 0.73 Cohen's kappa in comparison to gold-standard full-scalp EEG. Their results suggest that ear-EEG may be a real alternative to full-scale Polysomnography, especially for long-term monitoring.

Additionally, the relationship of changes in body temperature from ear-worn infrared thermometry and the circadian sleep-wake rhythm was initially shown based on a single subject [61] motivated by the possibility to measure body temperature at high accuracy on the ear (see subsubsection 3.1.3.1).

From a wearability perspective, sleep-related earables should place "rigid parts behind the ear and [...] soft materials at the concha and in the ear canal" [454].

3.1.2.3 Drowsiness. Drowsiness is the feeling of being abnormally sleepy. Drowsiness during the day affects 10-20% of the population and can have an adverse influence on physical and mental health, especially when operating vehicles or heavy machinery [479]. A summary table of drowsiness works can be found in appendix A.7. Bleichner and Debener [57] initially found that alpha power measured around the ear increased in the afternoon compared to recordings in the morning, indicating that the ear may be a suitable position for sensing drowsiness. Since then, drowsiness has also been detected by classifying light sleep onset with an accuracy of 80% when using in-ear EEG recorded in 20 minute sleep sessions [335]. Researchers have also explored multimodal sensing for microsleep detection. Pham et al. [374] designed a behind-the-ear device which collects data from EEG, EOG, EMG, and electrodermal activity sensors before streaming them to a mobile phone for classification. They demonstrated that microsleep can be detected on an unseen subject with average precision and recall of 76% and 85% respectively using a leave-one-subject-out cross-validation design with subjects suffering from sleep deprivation and narcolepsy [374]. More generally, Mikkelsen et al. [317] envision that earable drowsiness tracking methods may support long-term monitoring of daytime sleepiness disorders such as narcolepsy or hypersomnalence.

3.1.3 Thermal Regulation

One of the most important processes in humans is thermoregulation which refers to the ability to sustain a steady core body temperature (CBT) such as under different climate conditions or while exercising [488]. The ear can support measurements of body temperature and sweating, which will be described in the following.

3.1.3.1 Body Temperature. Under normal circumstances, body temperature ranges around 37.0°C $\pm$ 1°C [73]. In medical care, core body temperature is commonly measured in the ears as the tympanic membrane is located close to the carotid artery and, therefore, accurately reflects its temperature [62].

An infrared thermophile sensor pointing at the eardrum may be used for measuring body temperature according to medical standards. Bestbier and

Fourie [53] applied the principle to an earable form factor and achieved a small mean error of only 0.02 ± 0.52 °C. However, multiple works found that the principle requires per-user calibration because of ear canal shape differences and orientation of the sensor toward the tympanic membrane [53; 283; 300]. Alternatively, a thermistor measures surface skin temperature at the mastoid with high accuracy (0.03 °C mean error) [33]. According to Matthies et al. [304] the sensing principle requires an average heat-up time of 7 minutes and is easily influenced by environmental changes [205].

Across earable research, changes of body temperature in response to external weather conditions [61; 294; 40] and while exercising [61; 434; 294; 300; 91] were confirmed. This relationship enables various applications such as alerting or vital signs and parameter tracking based on the identified relationships. Additionally, the relationship between ear-recorded body temperature and ovulation (see subsubsection 3.1.5.5) as well as sleep (see subsubsection 3.1.2.2) has been shown.

3.1.3.2 Sweating. Commonly, sweating occurs in response to physical exertion, heat or psychophysiological arousal [421]. In general, the preauricular area has relatively high sweat gland density [426].

Matsumoto et al. [300] presented an earbud-type wearable prototype with a sweat rate sensor based on humidity sensing in the ear canal. They showed the relationship to physical exertion and amount of sweat based on a single user which they envision to apply for early detection and prevention of heat-strokes. Pham et al. [374] introduced a posterior ear device with integrated electrodermal activity (EDA) sensing. The normalised cross correlation between the ear EDA signal and a wrist-worn ground-truth evaluated with a single user was 0.37 [374]. As sweat gland activity is not symmetric between the two halves of the body it may be necessary to place electrodes on each ear to reliably capture sweating [377]. Overall, sweat sensing appears to be feasible on the ear but more research is necessary.

3.1.4 Mental State

A person's mental state entails, among others, emotions and stress, which trigger physiological changes that can be measured at the ear.

3.1.4.1 Emotion. Emotions include the feelings and thoughts of a person, which are commonly associated with their physical and psychophysiological state. In related literature, emotions are measured either dimensional (e.g., according to the valence-arousal model [406]) or categorical (e.g., happy, sad, angry, etc. [462]).

Based on in-ear EEG (see subsubsection 3.1.2.1 for the general feasibility), valence has been classified from low to high at 71.07% [35] to 94.1% [269] and arousal at 72.89% [35] accuracy (see appendix A.8). This accuracy is close to state-of-the-art performance of full-scale EEG [357]. Dimension-based scales can be applied to derive categorical emotion classes, which has been done based on in-ear EEG to predict happiness, sadness, calmness, fear at 53.72% accuracy [35]. Similarly, excited, relaxed, and negative emotions were predicted from in-ear EEG at 58.8% accuracy [269]. Interestingly, higher valence could also be associated with higher movement measured by an earable accelerometer [146]. Some models suggest up to eight basic emotions that can be observed in humans [380] which greatly over-exceeds what is currently possible to predict with data obtained from the ear.

Earable-based emotion tracking could enable monitoring patients and elderly remotely for more effective care-taking [35]. Additionally, facial expressions (e.g., facial action units) may be associated with emotions which are introduced separately in subsection 5.1.5.

3.1.4.2 Stress. Stress is triggered when an individual's mental or physical ability to cope with a situation effectively is over exceeded. It is commonly known that stress is associated with changes of heart rate and heart rate variability [441]. The previous subsubsection 3.1.1.1 gives an overview of the general feasibility of different heart rate sensing principles.

With earables, increasing heart rate and decreasing heart rate variability at the beginning of stress exposure elicited by a mathematical addition task was confirmed [283]. In a similar experiment, Suzuki et al. [439] also showed that higher heart rates at the ear are associated with stress using multivariate regression.

From the literature, the relationship between stress and skin impedance caused by sweat is well known [67]. Even though sweat sensing is possible with earables (see subsubsection 3.1.3.2), no earable works looked into the relationship between EDA measured at the ear and stress.

3.1.5 Health

The principle feasibility of the physiological parameters introduced in the previous sections can be applied to identify an individual's health status directly. Meanwhile, the location and phenomena of the ear also make it possible to derive insights into more specific bodily occurrences such as bruxism, coughing, sleep apnea, seizures, and even the reproductive system.

3.1.5.1 Bruxism. Bruxism is a movement disorder that is characterized by grinding of teeth and clenching of the jaw [420]. Tooth damage and headaches are common symptoms associated with bruxism. An initial evaluation with a single user successfully investigated the possibility to sense teeth grind and jaw clenching using electromyography on the ear (EMG) [125]. As jaw movements are closely related with ear canal shape deformations [378], jaw clenching and teeth grinding may be sensed from inertial sensors in the ear canal [404; 63]. Bondareva et al. [63] concluded that gyroscope-based sensing outperforms accelerometers when mixing with in-the-wild activities - achieving 76% and 74% accuracy for jaw clenching and teeth grinding, respectively.

3.1.5.2 Face Touching. Earables have also been used for detecting unconscious face touching which increases the risk of passing and spreading pathogens into the body, especially pertinent given the COVID-19 pandemic [221; 401]. Kakaraparthi et al. [221] used a hybrid sensing approach of thermal sensors embedded into an earable combined with facial skin impedance to monitor the user's face touching behavior. Using a deep learning model to combine the two signals resulted in an F1 score of 84.4% for touch detection and 70.1% for touch zone identification (rising to 90.1% for personalised models). Rojas et al. [401] introduced *Saving Face* which uses unmodified commercial earphones to sense the face touches through the distortion patterns in an ultrasound signal. The user's earphones are transformed into a sonar system with one earbud (positioned on the collar) emits the ultrasound signal which is captured by the microphone. The system was evaluated in a number of activities, achieving a sensitivity of 93.7% and precision of 91.5%.

3.1.5.3 Sleep Apnea. The interruption of breathing for 10 seconds or more during sleep is referred to as sleep apnea which, if not treated, poses a serious health risk [138]. As breathing results in subtle changes of the pulse wave in the PPG signal (see subsubsection 3.1.1.1), the interruption of breathing can be sensed with earables. In a whole-night study, 94.6% sensitivity and 93.4 % specificity was achieved for three out of six patients based on a PPG sensor worn at the tragus [469].

3.1.5.4 Epilepsy. As described in subsubsection 3.1.2.1, the brain activity of a person can be well quantified from ear-worn EEG. Epilepsy often entails seizures which are episodic abnormal neural activities that can result in body shaking or awareness loss. Given the high portability and unobtrusiveness of ear EEG, a natural application is the detection of seizures from ear-worn EEG for better health management and intervention [165; 219]. Bleichner and Debener [57] initially presented a periauricular EEG setup revealing epileptiform brain activity of a patient. Gu et al. [165] applied EEG posterior to the ear and reported 94.5% sensitivity and a false detection rate of 0.52 per hour to detect seizures of patients with focal epilepsy. Juez et al. [219] identified inter-ictal spikes in the EEG trace of a pre-diagnosed patient wearing an in-ear device. These spikes are known to occur in epilepsy patients.

3.1.5.5 Ovulation. Reproduction is critical to sustain the human species. The possibility to capture body temperature at the ear (subsubsection 3.1.3.1) and its relationship to ovulation creates a compelling use case for the earable platform. During the menstrual cycle, ovulation can be associated with the highest fertility, making it an important event to track when seeking pregnancy.

It is well known that ovulation can be associated with basal body temperature changes [42]. In an earable context, Luo et al. [284] recorded body temperature every 5 minutes during the night using a thermistor placed in the ear canal with 34 study participants. They could correctly predict ovulation within three days at 82.35% accuracy when tracking multiple cycles.

3.2 Respiration Tracking with Inertial Sensing

For oxygen gas exchange, the human chest rhythmically expands to perform inspiration (breathing in), followed by passive relaxation of the chest wall for expiration (breathing out) [74]. This process is referred to as *respiration*. Two key characteristics of respiration are inhaled and exhaled volume of air as well as rate. The number of breath cycles per time interval is commonly reported as cycles per minute (CPM). From medical research, it is known that particularly high or low breathing rates (< 8 CPM, > 24 CPM) can be associated with underlying health problems [109]

When tracking respiration rates in day-to-day scenarios, respiratory inductance plethysmography is the current state-of-the-art. A belt straps around the user's chest and abdominal wall to measure the expansion while breathing in and out [491]. Such devices are specialized and expensive equipment and not suitable for everyday use. A different method, applied for instance in sleep labs and under medical conditions, uses nasal cannulas made from plastic tubes which redirect the airflow to pressure transducers [347]. These tubes are uncomfortable to wear as they are placed inside the nostrils and are unhygienic when used multiple times or with different users. As an alternative, the dissertation proposes to leverage the inertial measurement unit (IMU) embedded in standard in-ear earphones for respiration rate tracking. This technology is potentially accessible to a broad set of users, as already today, earphones with integrated IMUs are commercially available (like e.g., the *Apple AirPods*).

Embedding respiratory sensing into headphones opens up a set of use cases where auditory feedback couples to breathing. For example, Harris et al. [177] suggest that auditory biofeedback can enable the control of the respiration rate of users , which can help with stress management or support guided meditations [418]. Other reasons to obtain respiration from the ear include the detection of interruption of breathing during sleep (see subsubsection 3.1.5.3) [469; 450] and more generally monitoring the vital respiratory state of the user for alerting [470; 53] or of workers in hazardous environments by embedding the sensor in hearing protection for faster intervention [296].

The following sections first introduce the necessary background and related work on wearable respiration sensing in subsection 3.2.1. Then, the working principle of an earable respiration rate sensing system based on IMU data and the required data processing pipeline is introduced in subsection 3.2.2. In subsection 3.2.3, the system is evaluated in a lab study with twelve participants to compare between standing, sitting, and lying on the back (supine). The system achieves 2.62 and 2.55 CPM mean absolute error (MAE) which is close to the accuracy of $\pm$ 2 CPM required for non-vital signs monitoring [270].

In sum, the dissertation contributes an evaluation of inertial-based sensing to measure the respiration rate of a user. The presented work was published as a conference paper at EarComp 2019 [448].

3.2.1 Background and Related Work

Previous research proposes a broad set of alternatives to the formerly mentioned state-of-the-art for respiration rate tracking.

Systems based on UWB [471], WiFi [1], or vision [297] have significant advantages because they are not attached to the user. However, they require specialized setups which can be complicated or might have problems with other people present in the room. Further techniques which do not require nasal cannulas to measure respiration from human breath include gas sensors that measure, e.g., volatile organic compounds [449] or humidity [325]. They still require to be placed in the air stream or have to be attached close to the area around mouth and nose.

Respiration rate, so far, was quantified by earables based on audio-, and heart beat-based sensing which will be introduced in the following (see appendix A.4 for details). As breathing is a biomechanical process, it produces tiny body movements and friction-induced sounds when the air enters and leaves the lungs. A microphone embedded in the ear canal could sense sounds propagating through the body during the respiratory cycle at 2.7 CPM MAE [296].

It was reported that such acoustic respiration rate measurements were reliable above approximately 12 CPM [160]. In [182], an accelerometer-equipped device wraps around the user's ear to measure respiration. They only evaluate the device with a single participant and do not consider gyroscope data as well as an in-ear form factor. Subtle respiration-induced changes in the cardiac cycle may be measured on the ear to indirectly derive respiration rate through PPG amplitude changes, blood oxygen variations, or the respiratory sinus arrhythmia [469; 53; 470]. From earable heart beat PPG signals, respiration rate was predicted at -0.558 $\pm$ 1.406 CPM mean error at the ear canal at rest [53] and around 3 CPM error in the ear canal, concha, and posterior auricular under varying motion activities [137]. With increasing motion, the performance of PPG-based respiration rate estimation decreased [137]. PPG-based sensing then could also be used to classify the more granular inspiration, and expiration phases at 81.5% sensitivity and 86% specificity [470].

The performance differences across sensing principles are relatively small. The U.S. Food and Drug Administration (FDA) acceptance criterion for non-vital signs respiration monitors requires a maximum error of $\pm$ 2 CPM [270] (or 10-20% of typical human CPM frequency). While sound-based respiration rate prediction with earables can not achieve such accuracy, PPG-based respiration rate estimation appears to be acceptable when the user is at rest.

Acceleration and gyroscope data have significant advantages in terms of cost and unobtrusiveness because many modern devices already come equipped with these inexpensive sensors (< 5$, single quantity). The idea to use the IMU's data is not new, and others have shown that the underlying principle does work. It has been implemented using chest belts [44] or smartwatches [172; 436]. Another approach which is particularly relevant because it is head-worn uses *Google Glass* smart glasses [188].

3.2.2 Working Principle

The following sections introduce the hardware setup to obtain inertial respiration sensing data and the working principle of the processing pipeline to calculate respiration rates from accelerometer and gyroscope signals.

Figure 3.2.: The Nokia Bell Labs eSense earphones [235] connect via BLE to transfer gyroscope and accelerometer data to a smartphone app.

3.2.2.1 System Design. The system leverages the *eSense* platform [235], which was kindly provided by *Nokia Bell Labs*. It comes equipped with a six-axis IMU (accelerometer and gyroscope) in its left earbud and connects via Bluetooth Low Energy (BLE). For respiration sensing, the x, y and z angular velocity in deg/s using the built-in gyroscope, and the x, y and z acceleration in m/s^2 based on the accelerometer is recorded (see Figure 3.2). Data is sampled at the maximum frequency of 50 Hz and without using the internal lowpass filtering functionalities of the eSense earbuds.

For recording the data, a mobile application in Swift for iOS was implemented that connects to the eSense earbuds. The data is stored locally on the phone, and timestamps are taken on a rolling basis as the Bluetooth packages arrive. To transfer the data to a computer, they are exported to a CSV file.

3.2.2.2 Data Processing Pipeline. To compute the respiration rate, the same processing steps are applied independently on gyroscope and accelerometer data. The approach expands upon the algorithm proposed in [188]. Steps (1), (2), (4), and (5) are applied to each axis. Additionally, steps (3) and (5) remove motion sensitivity.

1. To remove signal shifts and trends, a moving average window of 3 samples is subtracted from each dimension. Additionally, an averaging filter with a window size of 2 seconds is applied, corresponding to one respiration cycle at the maximum breathing rate (30 breaths per minute).

2. To inflate the data and create equidistant samples, a cubic spline interpolation is applied and the the signal is resampled at 256 Hz to account for small variations in timestamps due to Bluetooth latency.

3. Similar to Sun et al. [436], windows are discarded if movement is too high. If 3% or more of all accelerometer data points are above a threshold of 10 m/s^2 the entire sequence is excluded. Additionally, hard thresholding is applied for samples $\pm$2SD.

4. A bandpass Butterworth filter of order four and cut-off frequency of 0.1 Hz and 0.5 Hz is applied to remove noise. The selected frequencies are equivalent to 6 to 30 breath cycles per minute (CPM).

5. To further smoothen the signals while retaining the peak positions, a triangle filter with a width of 2 seconds is applied [176].

6. To make the results independent from changes on different axes for different postures, principal component analysis (PCA) is applied.

7. A spectral analysis of each principal component is performed using a Fast Fourier Transformation (FFT) with zero-padding. The maximum peak with the highest magnitude is assumed to be the respiration rate that is converted to CPM.

Figure 3.3 displays the raw signals captured from the accelerometer (left) and gyroscope (right) of one of the study participants sitting. The two graphs indicate how noisy the initial data signals are along all axes. As the user breathes, the accelerometer signal visibly oscillates around zero on the Y-axis and the gyroscope on the Z-axis. The second row compares a normalized ground truth signal captured using nasal cannulas hooked to a pressure transducer (PRS) with the filtered signal. The bottom row displays the three different spectra computed from the respective sensor signals. A red star indicates the maximum in each spectrum, which indicates a small error between the three different sensing principles.

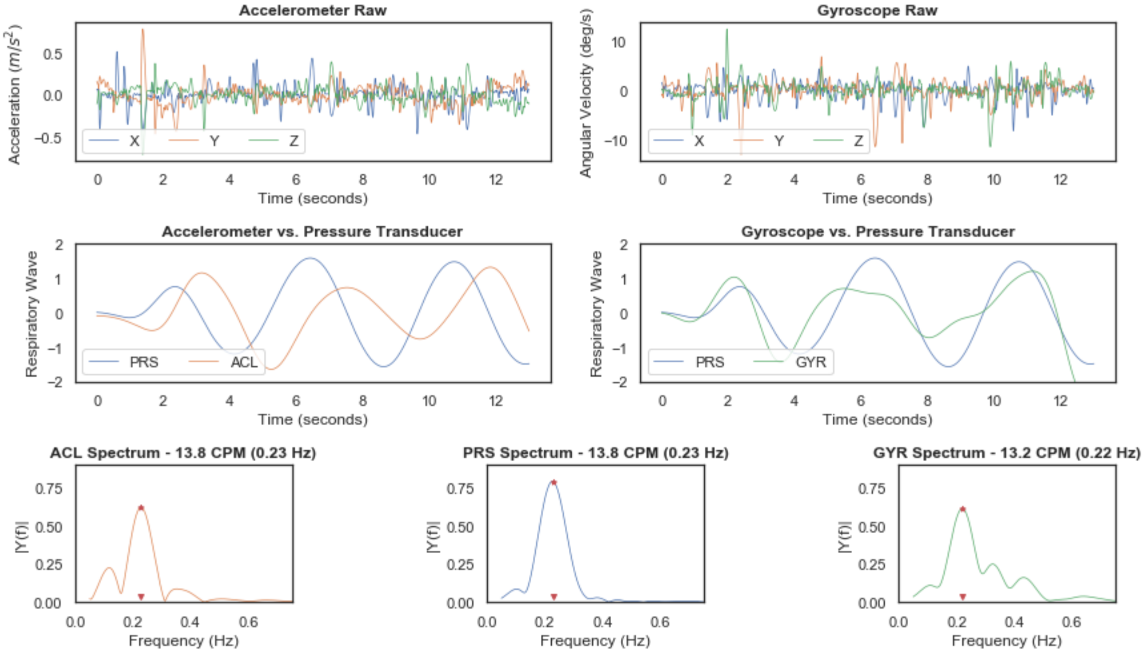


Figure 3.3.: Raw accelerometer and gyroscope data (first row). Filtered signals comapred to ground truth (second row). Spectrum of the processed signals with peak indicating respiration rate.

3.2.3 Evaluation

To evaluate the proposed system, the following sections first introduce the study design which is followed by the results achieved by inertial ear-based respiration rate sensing.

3.2.3.1 Study Design. Twelve participants were recruited (two female, ten male) between the ages of 21 and 39 (mean age 26) for a lab study. The mean height was 179 cm and weight 81 kg. Participants received no reward for their participation. The experiment was conducted in a room with a couch to lay down on and a stable chair with armrests for sitting down. Both *eSense* earbuds [235] were placed into the participant's ears (left earbud equipped with the IMU). Participants were hooked up to nasal cannulas connected to a pressure transducer as ground truth. This was a custom made monitoring device: a *RedBear BLE Nano v2* was wired to a pressure transducer that connects to nasal cannulas (see Figure 3.4). The device samples pressure data at a frequency of 50 Hz. The gold-standard pressure signal is filtered with the same data processing pipeline as described in 3.2.2.2, except no PCA is required. It also connects to the mobile application. No audio was played during the study.

Figure 3.4.: Participant wearing nasal cannulas (red, left) and the eSense earbuds (blue, right).

During the first phase of the evaluation, each participant was asked to breathe normally for one minute each in three different postures (standing, sitting, and lying) while otherwise keeping as still as possible. This step was followed by a second phase, in which participants were asked to perform a short 30-second jumping jacks session to increase breathing rate and produce a more dynamic dataset. After performing the activity, one minute of additional respiration data was recorded in each of the three postures.

Williams design generalized Latin squares [490] were applied to balance for first-order carryover effects introduced by potentially unnatural breathing behavior when asked to breathe on the spot. For the three different postures, this resulted in six different combinations, which were assigned to participants in a round-robin fashion. The order of postures for the first and second phase of the evaluation was identical. After completing all respiration tasks, participants were asked to fill in a short questionnaire, which included demographic questions (sex, age, weight, height) as well as a question inquiring whether participants felt that they had breathed naturally and space for free-text.

3.2.3.2 Results. In total, 72 minutes of breathing data were collected. A 20 s sliding window at a step size of 5 seconds is shifted over every one-minute data frame. This process yields 669 twenty-second breathing sequences. After removing the ones with too many movement artifacts, 253 remain. Ground truth respiration rates range from 7.6 to 22 in cycles-per-minute (CPM) for those sequences. To evaluate the agreement between the presented approach and the ground truth measurement, the Bland-Altman plots in Figure 3.5 were generated. In most settings, the observed differences are centered around zero and show no significant bias, also observable from the displayed mean error. The plots also show the limits of agreement (interval between +1.96SD and -1.96SD) that contain 95% of the measured differences. Additionally, Table 3.3 shows several performance metric. The results are broken down by body posture in Table 3.4. Overall, the performance of gyroscope is similar to the accelerometer but varies between postures. The best results are achieved by the accelerometer in the supine position, followed by similar results for sitting and standing. For the gyroscope, the results are comparable for all three postures.

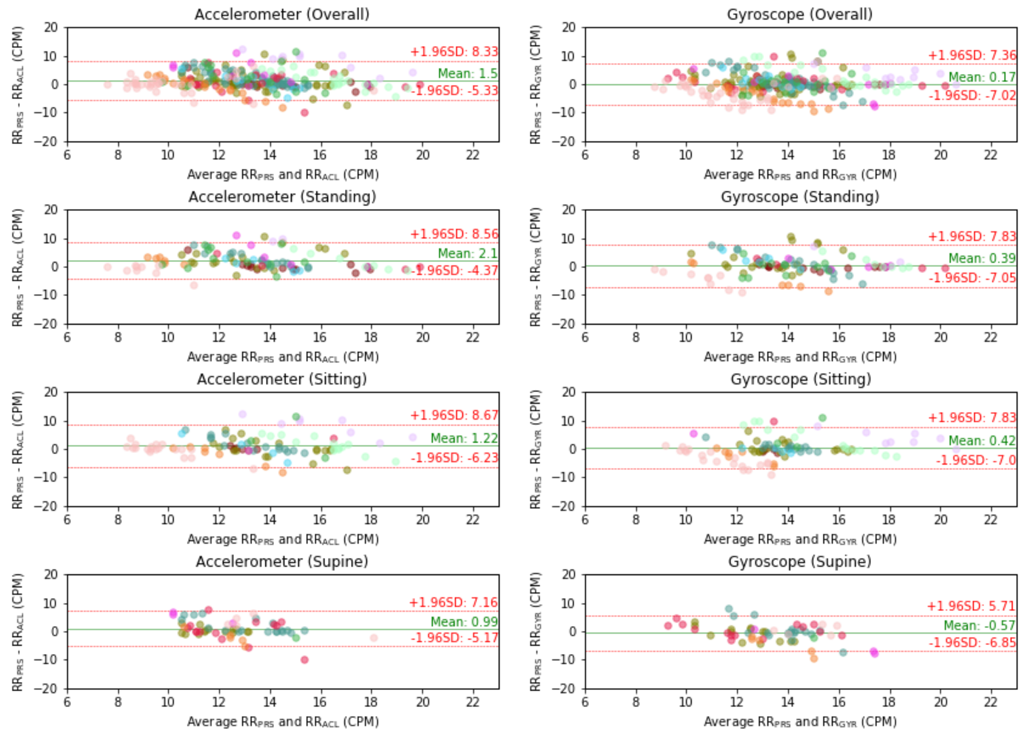


Figure 3.5.: Bland-Altman plots for the accelerometer (left) and gyroscope (right) according to the overall performance and for the postures standing, sitting and lying on the back. Per-user results indicated by color.

Table 3.3.: System performance in cycles-per-minute (CPM) for accelerometer and gyroscope across all three postures.

Sensor	**MAE**	**SD**	**RMSE**
Accelerometer	2.62	2.74	3.79
Gyroscope	2.55	2.63	3.67

Table 3.4.: Comparison between modalities and postures (MAE / SD) in cycles-per-minute (CPM).

Sensor	**Standing**	**Sitting**	**Supine**
Accelerometer	3.15 / 2.74	3.10 / 2.80	2.56 / 2.19
Gyroscope	2.45 / 2.22	2.74 / 2.64	2.68 / 2.00

3.2.4 Discussion

Generally, the proposed method is highly sensitive to motion artifacts. The first row in Figure 3.6 shows that better results are possible when setting the motion threshold to 1% and limiting the dataset to non-aroused participants (before jumping jacks). Introducing this limitation reduces the MAE to 2.09 CPM for the accelerometer and to 1.90 CPM for the gyroscope. Additionally, the results vary significantly between participants.

Figure 3.7 shows that even after raising the motion threshold to 5% participant P1 has much better results than P8 across a large range of breathing rates, especially for the accelerometer (MAE 1.21 ACC / 1.45 GYR for P1 vs. MAE 8.97 ACC / 4.58 GYR for P8).

The reason for the differences between participants are not entirely clear; however, bad fitting of the earplugs or differences in pose and anatomy could be a possible root causes.

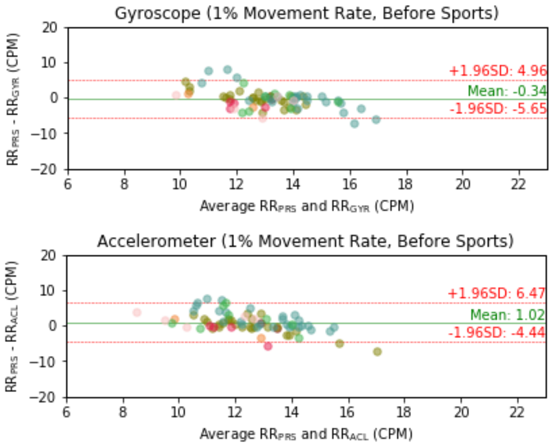


Figure 3.6.: Reducing the movement threshold increases the accuracy. Per-User results indicated by different colors.

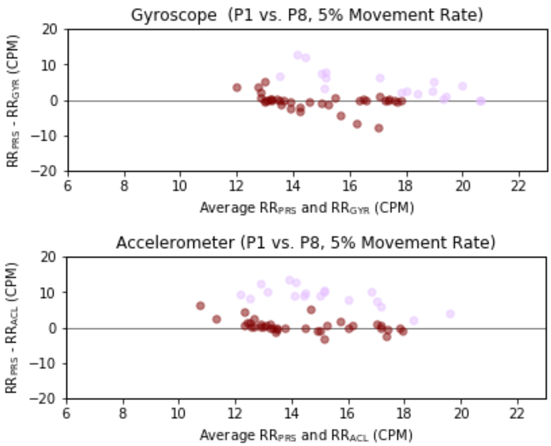


Figure 3.7.: P1 (dark) achieves a much lower mean error than P8 even at a higher movement threshold of 5%.

3.2.4.1 Comparison to Related Work. Table 3.5 compares the results to other related works based on inertial data on different locations of the body. Hernandez et al. [188] have evaluated smart glasses and also smartwatches [189] for the same three postures. Compared to [188] there are no significant differences between gyroscope and accelerometer results. Overall, the head seems to be a less suitable position for tracking respiration rates than e.g., the wrist. The presented results have higher error than what Hernandez et al. [188]; Sun et al. [436] reported, especially for the gyroscope. The root causes for the differences between earables and glasses is unclear, however, one possible factor would be a larger amplitude of motion when the sensor is positioned in front of the head compared to the ear which is more in the center of motion. According to Leong et al. [268], the thoracic spine moves back and forth, whereas the spine moves up and down while breathing. The resulting motion of the head could be different when measuring above the eye compared to the ear in different poses and on varying head positions. Another theory may be, that the differences could be caused by a dampening effect of the ear plugs' flexible caps which might absorb motion.

Table 3.5.: Performance (MAE / SD) in cycles-per-minute (CPM) of the system compared to related work.

Sensor	In-Ear	Glasses [188]	Watch [189]
Accelerometer	2.62 / 2.74	2.29 / 3.43	0.92 / 2.20
Gyroscope	2.55 / 2.63	1.39 / 2.27	0.38 / 1.19

Comparing the different poses, the results are similar to Hernandez et al. [188] for the accelerometer in the standing posture; however, the performance is worse for sitting and lying down. Additionally, the gyroscope's performance for standing is comparable, but the findings in this dissertation yield higher errors for the sitting and supine position.

3.2.4.2 Limitations. For the evaluation the nasal respiration was measured using a pressure transducer. After performing the physiologically straining task of jumping jacks, several participants reported the urge to breathe through the mouth afterward. Participants were not instructed to nose breathing before the study, but several participants reported that they "felt forced to not breathe through the mouth" (P2). An FDA-cleared chest belt based on respiratory inductance plethysmography is a more suitable methodology, which could support a more natural breathing behavior and therefore positively affect results. In addition, the experimental environment left things to explore visually (e.g., posters), which may have resulted in unwanted motion artifacts.

Since the presented work has been published as a paper, Ahmed et al. [4] later improved the performance by automatically selecting the best sensor channels and filtering out windows with too much motion which reduced the MAE to 1.62 CPM which is well below the required accuracy for non-vital signs monitoring.

3.2.5 Conclusion

Through this work, it was explored how in-ear inertial sensing may be applied for tracking respiratory rates. A data processing pipeline that combines multiple ideas from previous work was fit to the in-ear sensing use case. An evaluation compared accelerometer and gyroscope data for the three different postures standing, sitting and lying on the back (supine). The approach was validated by comparing measurements with ground truth data from nasal cannulas connected to a pressure transducer. In general, the results suggest that the ear is a less suitable position for measuring respiratory rates than, e.g., on the wrist. Overall, the solution has a high sensitivity to small motion artifacts. Nevertheless, it achieves stable performances for a subpopulation of participants that is suitable for non-vital signs monitoring (approx. 2 CPM absolute error).

To explore the inaccuracies between subjects, the potentially loose attachment of in-ear earbuds may be further investigated. Additionally, a possible dampening effect created by the eartip cushions may be interesting to understand. Lastly, a more advanced approach should fuse accelerometer and gyroscope data and the microphone signals to maximize performance.

3.3 Coughing Detection with Inertial Sensing

Diagnosis of patients suffering from aerosol-transmissible diseases such as the corona virus provides significant advantages as patients can be isolated quickly to avoid further spreading [185]. Objectifying the physiological state of the user allows foreseeing acute respiratory illnesses [329], whereas coughing and fever are the symptoms with high predictive value of, e.g., influenza [158]. Even though cough is a clearly perceivable indicator, objectifying the health status of patients by detecting episodes of increased cough might be helpful, e.g., to encourage seeing a doctor. To broadly succeed in containing a disease's outbreak, a ubiquitous technology must be applied. Ear-worn devices with an inertial measurement unit may be suitable to monitor cough [235]. Some earables even have a body temperature sensor (e.g., *Cosinuss°*), implying a holistic solution to predict respiratory illness. Therefore, the dissertation hypothesizes that kinetic earables can discriminate cough and non-cough activity to ultimately achieve episodes of increased cough discrimination.

In a twelve subject study, voluntary weak and strong cough, and five non-cough activities (laughing, throat clearing, swallowing, talking, being quiet) were collected under various conditions (laying down, sitting still, sitting fidgety, standing, walking). During the activities, an in-ear worn sensor records acceleration and gyroscope data. In total, 4,200 activity samples were collected. A single step classification pipeline (0.77 overall accuracy) serves as the foundation for statistical analysis to achieve episodes of increased cough discrimination. As a digression, that data is also analyzed for pose classification which could enable faster cough episode prediction.

In sum, the contributes are: (i) an evaluation of kinetic earables for cough event detection and (ii) a statistical analysis for increased cough episode discrimination. The presented work was published at the Augmented Humans conference 2021 [452]

3.3.1 Background and Related Work

Related work thoroughly investigated activity recognition with head-worn sensors [171; 207; 361]. Previous work covering cough detection, especially in the medical field, mostly relied on audio features. Smith [427] gives a good overview of the work in this area, while Pressler et al. [390] provide an introduction in the required signal processing chain. The techniques achieve partially high recognition rates >99%; however, mostly in laboratory environments. Monge-Álvarez et al. [328] reported excellent recognition rates in real-world settings; however, they do not apply leave-one-subject-out cross validation. Drugman et al. [126] showed that accelerometers can provide useful information for cough detection under various conditions and close to the signals of a microphone. Additionally, their work reveals which activities (e.g., laughing) are most easily confused with coughing. Therefore, the dissertation seeks to collect similar activity data samples. Mathie et al. [299] compared positions of accelerometers on the thorax for various breathing conditions with promising results. So far, kinetic earables have been used for tracking the respiration rate of the user (see the previous section 3.2), but not for cough event classification and episodes of increased cough discrimination. Hence, the dissertation researches the feasibility of such an approach.

Cough describes the voluntary or involuntary rapid expulsion of air from the lungs to clear the airways. A variety of severe viral infections and diseases such as COVID-19, influenza, or COPD are accompanied by cough symptoms [158; 324]. Detecting respiratory illness from cough allows quicker isolation [452] and monitoring disease activity [341]. The sounds and motion of cough create cough detection opportunities by sensing at the ears.

3.3.2 Evaluation

The following sections first introduce the data sampling soft- and hardware, which is followed by a description of the instructions that were given to participants for the self-guided data collection of different activity samples.

3.3.2.1 Activity Sampling App and Hardware. Because of the ongoing COVID-19 pandemic, it was not possible to collect data at the lab. Therefore, an app was developed that leads the user through the study fully autonomous at their home. It connects to the *eSense* hardware which was kindly provided by Nokia Bell Labs. It comes equipped with a 6-axis inertial measurement unit (accelerometer and gyroscope) in the left earbud sampling at 50Hz (see Kawsar et al. [235] for device details). It connects to the user's smartphone using Bluetooth Low Energy. The app uses *Flutter* with the eSense library [39]. For every activity (e.g., coughing) under each condition (e.g., standing), the app shows a countdown and then a full-screen call-to-action to tell the user to perform the activity. Users can pause at any time. The app stores all recordings and offers email export. As the approach should preserve privacy and require little data and compute, audio signals are not gathered.

3.3.2.2 Instructions and Participants. The app and earables are handed out to the participants and a phone holder is provided to place it within the line of sight. The base conditions *standing* and *sitting* should be performed on a stable surface and chair, respectively. *Fidgety sitting* involves shaking the body rhythmically and *walking* should be in a straight line. *Lying down* should be on the back. Activities to be performed are: voluntary weak cough and strong cough, laughing, throat-clearing, swallowing, talking and quiet. Every activity (e.g., cough) is performed ten times in a five second window under each base condition (e.g., sitting). Activities are shuffled randomly within each base condition. Users are encouraged to vary the duration and starting point of activities within the allotted time frame. All participants had instant messaging available to contact the study supervisor if something was unclear.

For the study, twelve healthy adult participants (8 male, 4 female) were recruited through a sample of convenience from the lab. The mean age was 25.3 years, the mean height 176 cm, and the mean weight 75.8 kg. In total, 1,200 cough, and 3,000 non-cough activities were recorded.

Figure 3.8 shows the different activities performed by one subject in the standing pose. Weak and strong cough signals are significantly different compared to the other activities. Especially strong cough shows acceleration on the x-axes and rotational forces on the y-axes, induced by the rhythmic motion of the upper body. Similar patterns are observable for laughing, however, at higher frequencies. Overall, throat clearing, swallowing, talking, and being quiet all lead to little motion. Looking at similar activity data under the walking and fidgety sitting base condition obviously shows high background noise.

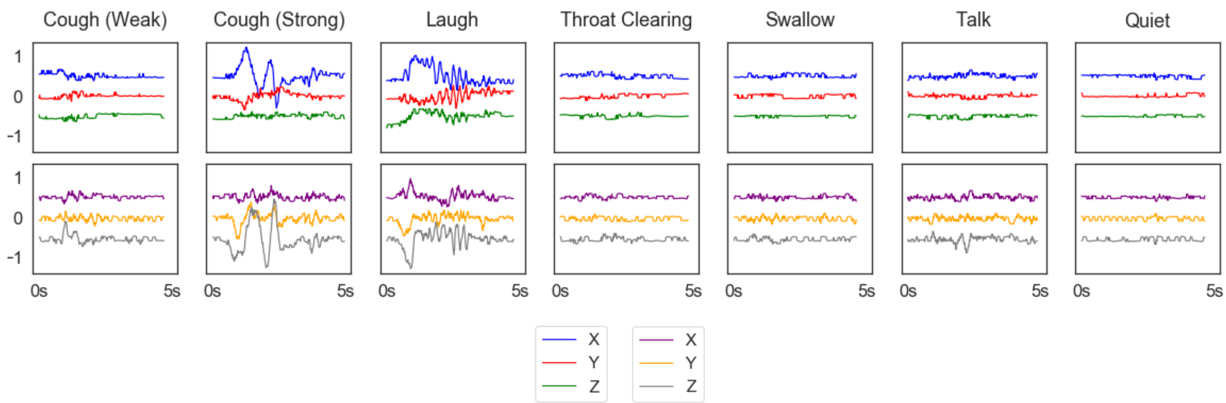


Figure 3.8.: Normalized acceleration (top) and gyroscope data (bottom) of the activities recorded while standing. Cough events show more motion than throat clearing, swallowing, talking or being quiet. Laughing has high motion energy but also higher frequency.

3.3.3 Classifier Pipeline

This section describes the pre-processing steps, the feature extraction, the training approach and the classifier results.

3.3.3.1 Pre-Processing. The data has to be pre-processed before being fed into the classifier pipeline. eSense does guarantee timeliness of data packages within a time-frame. Therefore, samples are evenly spread on the five second window and interpolation is performed to resample at 50 Hz. Windows which contain less than 200 samples are discarded. No further data cleaning is performed. No normalization or PCA is applied to retain activity signal energy and characteristic axial motions.

3.3.3.2 Features and Training. Automated feature extraction with *tsfresh* [103] is applied to all six axes. For training, cough and non-cough classes are balanced by randomly oversampling from all cough activities. XGBoost gradient-boosted decision trees is used in its default configuration [95], which outperformed a SVM and kNN classifier. A leave-one-subject-out cross validation strategy is applied. Each subject is held out for testing once and the remaining are used for training.

3.3.3.3 Results. Based on the collected simulated cough data data, relevant performance metrics were computed. Figure 3.9 shows the confusion matrix and Table 3.6 the performance metrics. The presented approach achieve 0.77 overall accuracy.

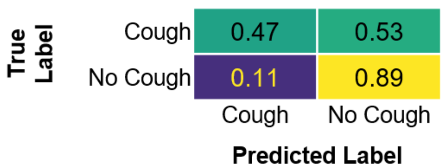


Figure 3.9.: Confusion matrix to predict cough and non-cough.

Table 3.6.: Performance for the two classes.

Class	**Precision**	**Recall**	**F1-Score**
Cough	0.63	0.47	0.54
No-Cough	0.81	0.89	0.85
Overall	*0.72*	*0.68*	*0.69*

3.3.4 Increased Cough Episode Detection

To evaluate episodes of increased cough activity, the exact number of coughs is not particularly important. Based on statistical analysis, a classifier better than random will be able to reliably discriminate individuals with more cough events than usual after a sufficient number of observations was collected. The previous sections introduced a simple classifier to differentiate cough from non-cough samples and, based on the confusion matrix of that predictor, in the following a statistical meta-analysis will be presented to illustrate how increased cough event detection can be achieved.

Statistically Discriminating Episodes of Cough. According to literature from the medical field, healthy adults cough up to 16 times per day [203]. For stable asthmatics, 282 coughs, and for chronic coughers, even 794 events per day are reported. Few coughs ($< 3\%$) occur during night time. The average sleep time of US adults is around eight hours per day [142]. If the 16 remaining hours are split into the five-second windows that were used in the study, this will yield a total of 11,520 observations per day.

A cough C as an event, can either happen in a given time-window or not. This condition makes modeling it as a Bernoulli variable a natural choice. From now on, for ease of understanding, the dissertation refers to episodes of increased cough as sickness (S). Assuming that coughing events in subsequent time-windows are independent given the factor that induces coughing (i.e., sickness), and assuming the probability of coughing given sickness $P(C \mid S)$ from the literature is the true value, the distribution of averaged-based estimates can be obtained that would be observed after periods of one day.

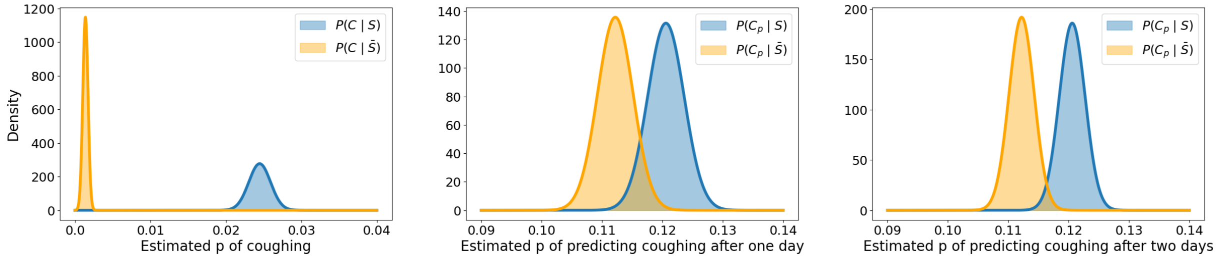


Figure 3.10.: The left graph displays the probability density function (pdf) of the average coughs for a sick (S) or non-sick (S) persons under perfect detection of coughs [203]. The middle and the right graph display the pdfs for the respective average predicted coughs of the classifier for one day (middle) or two days (right) of collected events respectively. Since the detection probabilities of the classifier biases the observed data, averages from the sick, and the non-sick distribution become more similar (left vs. middle and right). However, with more samples i.e., longer episodes (middle vs. right), the variance of the averages can be further decreased, which leads to less confusion of possible realizations of the two distributions. By observing the average number of coughing events of a user it can be decided if they are more likely to fall in either of the two distributions to decide if they are sick or not.

According to the central limit theorem (CLT), the average number of observed coughs $\hat{p}$ (for $P(C \mid S)$ or $P(C \mid S)$, would be distributed with $\mathcal{N}(p, N^{-1} \cdot p(1-p))$, where p denotes the true $P(C \mid S)$ or $P(C \mid S)$ respectively, and N is the number of samples used for estimation i.e. calculating the average.

When comparing the distribution of the estimates of the coughing probability given sickness (S) and non-sickness (S) for the number of events of $N = 11,520$ in a day, there is evidently little to no chance to confuse an estimate of $P(C \mid S)$ and $P(C \mid S)$ due to the clear separation between their distributions. In other words, it can be likely assessed which of both distributions generated the average of the data that is observed. The first graph in Figure 3.10 illustrates the clear separability of sick and not sick assuming 16 ($P(C \mid S) = 16/N$) coughs per day for healthy, and 282 coughs ($P(C \mid S) = 282/N$) for non-healthy individuals.

Unfortunately, coughs can not be observed directly but only the predicted coughs of the classifier (C_p) which may confuse cough events for non-cough events and vice versa. The confusion matrix in Figure 3.9 illustrates the probability of these misclassifications. Through these, it is only possible to estimate $P(C_p \mid S)$ via averages of predicted coughs, but not $P(C \mid S)$. Naturally, $P(C_p \mid S)$ and $P(C_P \mid S)$ are biased (through the confusion) and will not yield the same values as $P(C \mid S)$ and $P(C \mid S)$. However, the distribution of both estimates may still be different enough to decide whether a healthy or a sick individual generated the observed average.

To assess the distribution of these averages, the probability of $P(C_p \mid S)$ is calculated, which can be readily derived as

$$P(C_p \mid S) = P(C_p \mid C, S) \cdot P(C \mid S) + P(C_p \mid C, S) \cdot P(C \mid S) \tag{3.1}$$

$$= P(C_p \mid C) \cdot P(C \mid S) + P(C_p \mid C) \cdot P(C \mid S). \tag{3.2}$$

Note that $P(C \mid S)$ and $P(C \mid S)$ are the probabilities given above, while $P(C_p \mid C)$ and $P(C_p, \mid C)$ can be calculated from the confusion matrix in Figure 3.9. Further note that from Equation 3.1 to Equation 3.2 it assumed that the probability of predicting a cough is conditionally independent of the health status of the person, i.e., the confusion of the classes does not depend on if the person is sick or not, but merely on whether a cough was observed. The probability of $P(C_p \mid S)$ can be calculated analogously.

As before, the distributions of the estimates of $P(C_p \mid S)$ and $P(C_p \mid S)$ can be derived through the CLT. The resulting distributions of the (predicted) average coughs for one and two days can be seen in Figure 3.10. Note that for two days ($N = 2 \times 11,520$), the distributions become more narrow, as the variances get smaller for higher N.

This information can now be utilized to classify episodes of increased cough based on the average number of observed coughing events. One possibility is to decide for sickness and non-sickness based on if the average is more likely to fall in either the distribution of the cough estimates given sickness or non-sickness. Alternatively, the distribution of the average coughs for a healthy person can be used to classify them as sick if the probability of the observed average is below a certain threshold i.e., confidence level. While the former approach weighs in misclassifications for sickness equally, the latter tries to avoid false positives. Generally, decision rules can be derived based on costs associated with each type of error.

Note that for the derivation of this result it was assumed the true probabilities are known. Unfortunately, this information would seldom be given in practice but also has to be estimated. Here, longer episodes (e.g., a week) of sickness or non sickness could be used to achieve estimates closely reflecting the true value. For example, a system could observe an individual in a known healthy state to compute the personalized $P(C \mid S)$ over time. Then, significant changes in the average number of predicted coughs could be detected to identify sickness on a more personalized level.

3.3.5 Discussion

Since publication of the results presented in this dissertation, other authors have explored cough detection with earable sensing data. Zhang et al. [503] introduced an algorithm based on template matching that allows tuning sensitivity and specificity to the desired use case with about similar performance to what was presented in this work under noise-predicated situations. Interestingly, Doddabasappla and Vyas [123] could show that the ear had the best performance to predict simulated cough from accelerometer data in comparison to the chest, stomach, shirt pocket, and upper arm [123; 124]. Another work could show that combining audio and motion data obtained at the ear decreased the false positive rate of cough detection and achieved 83% sensitivity, and 91.7% specificity, which is an increase by 55% compared to audio-only [341]. An advantage of audio-based cough prediction is the availability of audio cough datasets, even for specific diseases such as COVID-19 [343].

Reliably classifying pose and, in general, if the user is at rest could improve the cough detection as the samples would be less noisy. For example, evaluating the presented cough classifier for the sit, laying down, and stand still poses increases the possible accuracy by 9%. Therefore, an advanced pipeline may detect situations where the user is generally at rest and then quantify cough which could reduce the required monitoring time.

It remains to be evaluated how proposed earable-based techniques perform based on real-world cough sounds and motion data from the field.

Digression: Pose from Kinetic Earables. Detecting simple activities such as walking with head-worn sensors has been done by others [171; 202]. As the collected dataset contains airway activity related noise the dissertation investigates the achievable performance for classifying poses under these conditions. A classifier that detects the different poses under varying activities was evaluated. Again, *tsfresh* [103] feature extraction is applied and an XGBoost [95] model is trained and evaluated with a leave-one-subject-out cross-validation strategy. Figure 3.11 shows the confusion matrix of the experiment.

Walking is detected reliably and mostly confused with fidgety sitting. Lying on the back can be well distinguished as the gravity vector changes. Sitting still and sitting fidgety as well as standing are confused most easily. Sitting and standing have no distinguishable characteristics. Fidgety sitting is subject to the kind of movement performed by the participant, which might be little. As poses usually change less frequently than within the allotted 5-second window that was recorded, further performance improvements are possible. A classifier that distinguishes between "at-rest" and "moving" (walk, sit fidgety) achieves satisfying results (see Table 3.7).

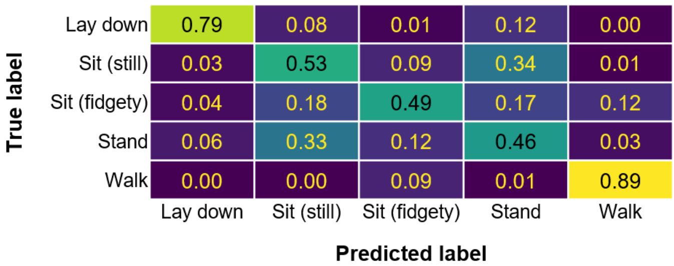


Figure 3.11.: Confusion matrix of different poses.

Table 3.7.: Detecting if the user is moving or at rest can be well derived from kinetic earables under interfering activities.

Class	**Precision**	**Recall**	**F1-Score**
Moving	0.87	0.75	0.81
At-Rest	0.85	0.93	0.89
Overall	*0.86*	*0.84*	*0.85*

Limitations. The conducted experiments assume that motions of simulated and real activities are similar, which may not stand in real-world settings, especially for cough. For example, the characteristics of coughing symptoms of COVID-19 are distinctive for the disease [254]. Additionally, it is assumed that the probability of predicting cough is conditionally independent of the health status of the person. There may be other activities that create confusion to classify coughs. In related work, audio signals were reported to have the most significant value to predict cough [126]. This work opted for a non-audio based solution due to privacy. Still, a microphone could easily be employed on the ear. In future work, collecting a dataset of sick people and in-situ over an extended period could show the proposed technique's robustness and applicability of the mostly theoretical assessments to discriminate cough from kinetic earables.

3.3.6 Conclusion

The presented study could show that increased episodes of cough discrimination can be achieved with kinetic earables based on a simple brute-force classifier and statistical analysis. The collected dataset from twelve subjects contains 4,200 activity samples which were used to perform a leave-one-subject-out cross-validation. Compared to an inaccurate one-off classifier, the presented approach benefits from modeling the probability of observing coughs. The sensing component relies on a single 6-axis IMU in the left ear and requires no microphone. The concept may be easily transferred to other head-worn devices such as smart glasses, for which the ear hook attaches in a similar location as the eSense earable. Future work may collect earable cough data over prolonged periods from sick and non-sick patients to verify the suggested principle.

4. Movement and Activity

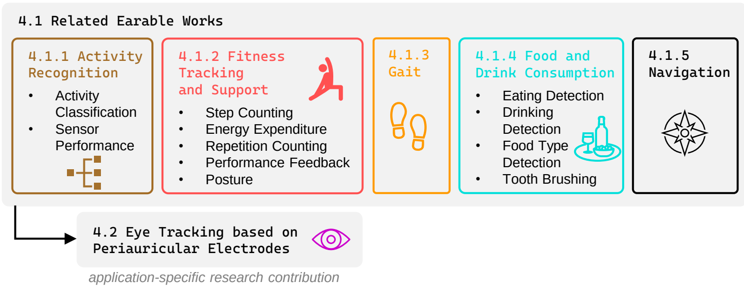


Figure 4.1.: Structure of the “Movement and Activity” section according to different application areas. The section "Eye Tracking" is separated out because of the sensing contributions and studies conducted in this dissertation.

This chapter focuses on movement and activity applications related to ear-based sensing. Figure 4.1 gives an overview of the structure of the chapter. Researchers have explored how the earable platform can be used to sense user movement and infer information about activities the user is performing. Movement detected at the ear can be classified into discrete classes to infer a user’s pose, how they are moving, or what type of activity they are undertaking. Beyond just classifying the sensor data, researchers have also explored how physical quantities can be derived from the user’s movement to provide useful information for a wide variety of applications including fitness tracking, gait analysis, food and drink consumption, inertial navigation. In addition, eye tracking is a popular modality for activity recognition which is explored in depth with earables in section 4.2.

4.1 Related Earable Works

The following sections are structured according to different application areas of movement and activity, as illustrated by Figure 4.1.

4.1.1 Activity Recognition

Sensor-based activity recognition has been a staple of the of the ubiquitous computing community for many years, both as a technical pursuit focused on improving recognition rates but also as a means to understand activity context that can be used to personalise devices and systems. Earables are compelling for activity recognition because many people wear ear-based devices, such as earphones or hearing aids, as they perform their everyday activities. In contrast, other wearables are commonly abandoned as activity trackers due to their inability to be incorporated into a person's everyday life (e.g., knee or waist trackers) [178; 36].

4.1.1.1 Activity Classification. Prior research has explored how ear-based devices can successfully recognise a wide variety of activities. For basic activity recognition tasks, such as determining whether a user is walking or running, earable-based accelerometers on their own provide an almost perfect accuracy of 99% [84]. Similarly, Min et al. [318] reported an F1-score of 95% when determining the mobility of a user (stationary, walking, stepping up, stepping down) and 80% when performing a head gesture (nodding, shaking) by utilising both the accelerometer and gyroscope and using a nearest-neighbour classifier. Hammour and Mandic [174] show that motion artefacts from in-ear EEG can be used to classify four basic activities (sitting, walking, speaking, and chewing) with 85% testing accuracy. For estimating a user's understanding of an online lecture, Kim et al. [243] classified whether a user is gazing at a monitor or looking down at the desk based on the accelerometer and gyroscope of the earable, achieving F1-scores of ≈ 0.92 and ≈ 0.90 respectively. These results demonstrate how earables can detect simple movements and gestures with high accuracy.

However, more advanced activity recognition tasks show more variable results. Atallah et al. [29] explored one-versus-all classification for ear-based accelerometer sensing using different activities groups based on the "compendium of physical activities" [7]. Activity recognition using a nearest neighbour classifier was similar for the high-level (running and cycling) and low-level (preparing and consuming food, socialising, reading, and getting dressed) activities with approximately 65-70% F1-score. However, medium level activities (walking and cleaning) performed much worse than both with approximately 50% F1-score. Similar results were also found in previous research by Atallah et al. [26], although Nirjon et al. [345] reported 96.8% accuracy when classifying only the level of activity (i.e., physical intensity) rather than the activity itself.

Advances in the underlying classifiers may also increase recognition performance on the ear. More recent work has explored how end-to-end deep learning can classify five scripted activities (nodding, speaking, eating, staying, and head shaking) with an F1-score of 82% [256] and physical exercises with 82% accuracy (squats, lunges with dumbbells, alternating bicep curls, sit-ups, push-ups, sitting overhead dumbbell triceps extensions, standing dumbbell rows, jumping jacks, sitting dumbbell shoulder press, and dumbbell lateral shoulder raises) [433]. Across these works, earables are able to achieve good recognition performance despite the difficulty in detecting movements of other parts of the body.

Researchers have addressed the difficulties of using earables to detect complex movements by combining them with other devices in a complementary manner, leading to greater recognition rates than are possible with any single device. Strömbäck et al. [433] recorded motion data from ten participants performing ten different exercises while carrying commodity wearable devices – a smartphone, smartwatch, and earbuds equipped with inertial sensors. Fusing the data using deep-learning techniques across wearables achieved 96% activity recognition accuracy, which was significantly higher compared with using the data from only a single wearable (earbuds - 82%, smartwatch - 94%, and smartphone - 85%) [433]. Radhakrishnan et al. [394] pair earables with dumbbells, both augmented with an accelerometer and a gyroscope, in order to classify the free-weight exercise a user is performing with a test-set accuracy of 96.85% from 3 exercises (bicep curls, triceps extension, and lateral raises), 93.72% from 6 exercises (also including squats, lunges, and side bend), and 88.6% from 12 exercises (also including seated barbell shoulder press, inclined chest flyes, weighted crunch, dumbbell triceps kickback, barbell deadlifts, and alternating bicep curls). Other work suggests fusing accelerometer and gyroscope data from an earable with optical flow from cameras can also improve classification accuracy for basic gestures including reaching for items and dressing oneself [307].

4.1.1.2 Sensor Performance. Accelerometers and gyroscopes are the main sensors used to detect motion of the earable for activity recognition, with the exception of one paper which explored the motion artefacts of in-ear EEG [174]. For many applications it is only the accelerometer that is used [29; 440]. Min et al. [318] found that the accelerometer significantly outperformed the gyroscope, and that fusing the two sensors resulted in a marginal performance increase over the accelerometer-only approach (only 1% for head gestures). Using only the accelerometer is compelling because they are more energy-efficient than gyroscopes [318]. However, despite finding the same relative performance differences between the three conditions, Laporte et al. [256] found the difference in performance between accelerometer-only and hybrid approach to be larger, and between accelerometer and gyroscope to be smaller, when using end-to-end deep learning with F1 scores of 75% (accelerometer), 69% (gyroscope), and 80% (both).

Comparisons between the ear and other locations show it to be among the best positions for activity detection. Atallah et al. [29] compared accelerometer placement of seven different locations on the body whilst tracking different high-level activities (including preparing food, getting dressed, cleaning, and socializing) of eleven users. The ear was second best after the knee, without significant loss in classification accuracy, and better than the chest, arm, wrist, waist, and ankle. More recent work has also shown similar results with only the knee and shin outperforming the head for activity classification for activities including climbing stairs, jumping, lying, standing, sitting, running/ jogging, and walking [440].

However, more recent work that uses deep learning suggests that other commodity wearables may outperform earables when classifying physical exercises. Earables achieved an accuracy of 82%, lower than both a smartwatch (94%) and smartphone (85%) [433]. Earables are also susceptible to falling out during heavy movements, even when adjusted correctly, and therefore certain activities may be less suitable for ear-based devices [195]. In addition, earables produce greater wearing variability in comparison to a smartwatch, with differences of roll and pitch of about 10-20 degrees in earables compared with 3-8 degrees [322] in smartwatches. Despite these disadvantages, earables provide good recognition performance across a wide range of activities, and when used in combination with other wearable sensors can further improve recognition accuracy [433].

4.1.2 Fitness Tracking and Support

Sedentary lifestyles account for 25% of all medical expenses and cause millions of deaths worldwide [15]. Beyond just classifying physical activities or sports, as shown in the previous sub-section, earables can help to improve and manage a user's fitness and health [40; 258; 387; 198; 65; 84; 29]. To date earable research has covered a range of different fitness activities ranging from weight-lifting [394] and general exercise routines [433; 330], to sports including cycling [65; 28; 29], rowing [28], climbing [358], and basketball [195].

4.1.2.1 Step Counting. Counting how many steps a user takes is one of the most commonly used metrics for tracking and assessing physical activity, and for encouraging users to stay active [43]. Earables are ideally suited for step counting because the body acts as a filter which stabilises the head during locomotion [164; 232]. Prakash et al. [387] show the initial feasibility of step counting at the ear with a variety of walking speeds including very slow, slow, normal, and running that can be tracked with 95% accuracy [387]. They also propose to detect and measure jumping to assess the physical health of a user, with a limited user trial reporting jump heights of within 1-3cm of the ground truth.

4.1.2.2 Energy Expenditure. Motion of the human body measured on the ear directly relates to physical activities performed by the user, and this relationship can be used to infer a user's energy expenditure [144]. This approach provides a more generalised way of measuring a user's activity in comparison to step counting. Energy expenditure can be estimated from acceleration forces measured at the ear with high correlation confirmed by calorimetry (r=0.92 [258], r=0.74 [65]). This relationship was confirmed for lying down, standing, computer work, vacuuming, walking stairs, slow walking, fast walking, slow running, fast running, cycling, and rowing [28; 65]. Consequently, a mean absolute deviation of only 27 kcal per day was achieved [65]. Beyond tracking a user's activity to monitor their fitness, energy expenditure prediction can be combined with dietary measures to maintain long term weight loss [258] or to monitor patients remotely [26]. Also, Nirjon et al. [345] used the concept of energy, derived from the accelerometer, to detect the activity level of a user which was then used to control music being played.

4.1.2.3 Repetition Counting. Strömbäck et al. [433] used earables to count the number of repetitions performed whilst exercising for a variety of ten physical exercises. Whereas accelerometers are favoured for activity classification, the authors found that gyroscope data results in better performance for repetition counting, although only marginally. Earables achieved a mean absolute repetition counting error of 1.31 (all results reported are for gyroscope), outperforming a smartphone (1.61), but significantly worse than a smartwatch (0.34). Interestingly though, the percentage of exercise sets within two repetitions of the actual number of repetitions was slightly higher for the smartphone (82.95%) than for the earable (81.33%). Again the smartwatch significantly outperformed both with 98.70% of exercise sets within 2 repetitions.

4.1.2.4 Performance Feedback. Earables can also be used to provide feedback to users whilst performing a physical activity in addition to analysing their form afterwards. Radhakrishnan et al. [394] not only classified the type of exercise a user was performing, but also developed a feedback system for use during weight-based exercises that combined earable motion data with motion data of a dumbbell which detected exercise mistakes (94%) of 33 users which helped to reduce subsequent mistakes by more than 10%. Similarly, Motokawa et al. [330] explored how acceleration data from the ear and chest can be combined to provide real-time corrective feedback during planks. Rather than providing real-time feedback, Pansiot et al. [358] used the accelerometer in an earable to collect data so that climbers could assess the fluidity, speed, strength-to-weight ratio, and endurance of their performances and to provide insights into their proficiency level. Hermann et al. [186] applied earables to support life-saving cardiopulmonary resuscitation (CPR). Based on an evaluation with twenty users on a test dummy, acceleration data measured at the ear of the rescuer can predict chest compression depth and rate at 5.9 mm and 1.6 cycles per minute median absolute deviation. Though that was worse than on the rescuer's wrist or chest, the performance of earables was still within the acceptable range for real-world feasibility of CPR.

4.1.2.5 Posture. Poor posture when sitting at a desk can cause back pain and poor circulation that can lead to other health issues [110; 180]. Earables equipped with an accelerometer and a gyroscope can be used for corrective posture feedback by detecting when a user leans forward with their head while sitting down [442; 395]. An initial simulated experiment based on five participants using their smartphone or laptop 30 minutes each yielded perfect precision at 89% recall to detect a forward-leaning posture [395].

4.1.3 Gait

Gait describes the motion performed by a user when walking. Earables can accurately detect several temporal and spatial gait parameters including heel contact time, toe off event, swing time, stance time, stride time, step cycle time, and step asymmetry [32; 210; 211]. The relationship between gait-specific motion data and acceleration measured at the ear has been confirmed in conjunction with gold standard procedures including an in-shoe pressure measurement system [210], a force-plate instrumented treadmill [32; 30], and a high-speed camera [211; 212]. Extracting these precise gait parameters provides clinicians with the opportunity to monitor rehabilitation from either surgery or a stroke, or progression of pathologies including osteoarthritis [32]. Ear-worn devices are a much cheaper and simpler platform for detecting gait patterns and parameters than traditional gait platforms and force plates, and have the added advantage that they can be deployed in-the-wild.

Gait parameters have been successfully extracted using earables and applied to assess health outcomes of an individual. appendix A.10 details all earable papers that have explored gait-related parameters. Recovery after surgery can be tracked using acceleration signals at the ear which capture the irregularity of a user's gait pattern [27]. Atallah et al. [27] used this possibility to both assess the recovery over time as well as approximating the user's gait to a healthy control group. Similarly, Jarchi et al. [213] conducted a validation study which showed that stride time, amplitude asymmetry, and step time measured at the ear improved one year post-surgery. The amplitude asymmetry level was also found to correlate with the Knee Injury and Osteoarthritis Outcome Score [210].

Impairment of gait, e.g., because of skeletal malfunctions or due to aging, can also be extracted directly using earable devices. Accuracy of more than 95% was achieved for predicting truncal [31] and lower limb impairment [25; 31] using an accelerometers on the ear. Lorenzi et al. [280] used an inertial measurement unit consisting of an accelerometer, gyroscope, and magnetometer to measure the freezing of gait of Parkinson's disease patients. The choice of an ear-based sensor was motivated through wearability and also the need to provide auditory feedback to feedback to the patient. The risk of falls in elderly people has been clinically measured using the Tinetti Gait and Balance Assessment (TGBA) [465]. King et al. [244] discovered that certain aspects of TGBA can be assessed base on motion data from the ear (2 out of 17 test-related activities), however this small fraction is likely not enough for practical use. Similarly, ear-based devices have been equipped with accelerometers for detecting fall events aimed at elderly patients [478].

4.1.4 Food and Drink Consumption

Dietary monitoring involves detecting both when someone consumes food or drink and, ideally, also what they are consuming. The ability to automatically monitor dietary activity can take the burden of self-reporting away from a user or patient and assist with mindful eating, tackling unhealthy behavior, and preventing diseases by supporting healthier diets [49; 13; 14]. Earables are suitable for automatic dietary monitoring because they are non-invasive [415; 360; 150; 366; 364], socially acceptable [415; 47; 56] and can be worn throughout the day [366; 346; 150]. appendix A.9 details all of the work that has explored how to detect when and what a user is eating or drinking on the earable platform.

4.1.4.1 Eating Detection. The first step of dietary monitoring is detection of eating events. The repetitive nature of jaw movement when chewing and eating is beneficial when trying to predict eating phases under real-world conditions [49] and can be detected using several sensing principles. Air- and bone-conduction microphones sense chewing sounds; proximity, piezoelectric, and inertial sensors track ear canal deformations induced by mouth motions; and electromyography (EMG) quantifies chewing-based muscle activity directly. On their own these different sensing modalities achieve roughly similar accuracy rates ranging from 80-90% in field experiments, however fusing multiple sensors can further improve recognition accuracy [281; 55; 359].

Each sensing principle used to detect eating behaviour has advantages and trade-offs. For eating detection, body-internal vibrations and sounds can be measured due to the ear cavity that are amplified by the ear's physiology [56; 13]. However, external and background sounds not relating to the chewing activity should be dampened for optimal recognition [359], which remains a critical challenge for audio-based approaches [150]. This issue is further exacerbated when chewing softer foods as the amplitude of the chewing signal is much lower compared with crunchier foods [281]. An additional microphone which measures and filters sounds from the surroundings can improve performance [310; 364], and deep learning approaches have also been shown to increase recognition accuracy up to 77-94% even with ambient noise [150]. Audio-based approaches are still an active research area because they possess a key advantage to other sensing modalities as microphones are already commonly embedded in many commercial ear-worn devices.

Motion-based approaches sense movement of the jaw when eating. They do not face the same privacy concerns and are more robust against soft food types than audio-based approaches [281]. Similar to audio-based approaches, motion-based eating detection also suffer from signal noise induced by unrelated body movement [281]. Bedri et al. [49] showed how it is possible to filter out undesired motions by using additional body-worn sensors, in the form of an IMU behind the user's neck, resulting in an F1-score of 80.1% based on field experiments. However, having the additional IMU in a form-factor that a user is likely to use is still an open challenge.

Proximity sensors can be used to detect ear canal deformation as a result of jaw movements when eating. They are compelling as they require less power and do not suffer from the same level of privacy concerns in comparison to audio-based approaches [47]. Bedri et al. [47] measured ear canal deformation using three orthogonal proximity sensors and a gyroscope embedded in an off-the-shelf earpiece which resulted in 95.3% accuracy when detecting eating events. Also, Bedri et al. [49] explored the use of an in-ear proximity sensor but found that users found it uncomfortable when wearing for prolonged periods and especially when eating.

Other sensing principles have been used to detect eating at the ear via jaw movements. Wet EMG electrodes located behind the ear on the mastoid detect the jaw's muscle activity [55], PPG senses the changes in blood flow as a result of the jaw movement [360; 359], and piezoelectric strain gauges located on the lower jaw directly measure jaw movement [415; 140]. EMG and PPG were both significantly outperformed by audio-based detection, yet when fused resulted in an overall increase in accuracy [55; 360].

Multiple lab studies investigated how to predict individual bites and chew strength on the ear using pressure, bend, and piezoelectric strain sensors [414; 200]. However, individual chews are more complicated to predict than general eating activity due to the higher temporal resolution that is required [359].

4.1.4.2 Drinking Detection. Staying hydrated is important for cognitive function and overall health [399]. Tracking the consumption of liquids (including liquid foods) is a challenging detection task for earables because it lacks the characteristic chewing information (e.g., [414; 281; 150; 364]). Some works reported good results based on sound sensing for a subset of study participants [424], however drinking detection can not be reliably detected with earables (see appendix A.9).

4.1.4.3 Food Type Prediction. In addition to detecting when the user eats, automatic dietary monitoring should detect what food the user is consuming, and how much of it they are eating. The amount and type of consumed food are key contributors to the success of weight loss maintenance [252]. Research to date has focused on distinguishing foods of different textures in a lab setting using audio data collected at the ear, with a success rate of 79% [366]. Alternatively, cameras located on the ear can be triggered by a microphone when chewing is detected and images are taken of the meal which can be later analysed for an overview of the user's food-intake [273]. No field studies or studies involving food type predictions of complex meals have been undertaken.

4.1.4.4 Tooth Brushing. Closely related to food and drink consumption is tooth brushing, as sugary diets in particular can lead to cavities and problems with dental hygiene. Researchers have explored how tooth brushing location can be tracked using earables in order to help form better tooth brushing habits [356; 389]. Prakash et al. [389] modified an off-the-shelf earphone speaker to detect the vibrations from tooth brushing with 89% accuracy over 7 locations of the upper and lower teeth. Ouyang et al. [356] explored a larger number of locations (16) and achieved a similar prediction accuracy by using a combination of throat and ear microphones.

4.1.5 Navigation

Navigation can be achieved by leveraging the inertial sensors in an earable to track the position and orientation of the user over space and time without having to rely on a GPS connection [493; 6]. Listening to music or wearing earphones is common when travelling and the stable attachment and fewer random movements of earables in comparison to other locations on the body makes them a suitable platform for inertial navigation [136; 157].

However, the accuracy of inertial navigation is dependent upon a fixed global reference point which is commonly the magnetic field of the earth and detected by a magnetometer. In earable devices, the close proximity of communication circuitry and the speaker introduce significant electromagnetic noise which has so far proved problematic for inertial navigation on the earable platform [136]. To overcome this problem, Ferlini et al. [136] developed an automatic magnetometer calibration method in combination with the user's smartphone that reduces the average error from 30 degrees to less than 5 degrees at any given time. Based on a 9-axis inertial measurement unit and the described magnetometer calibration method, an inertial navigation drift of 0.15m/s when using one earable and 0.11m/s when using two earables could be achieved [6]. Gong et al. [157] presented a deep learning based pipeline that fuses the inertial data from a smartphone and earables. This pipeline takes into account the reliability of each sensor at any given point in time prior to fusing, and only relies on the earables accelerometer and gyroscope readings for inertial navigation, achieving accuracy improvements in comparison to other state-of-the-art navigation models.

To navigate indoors, Schindler et al. [416] proposed an ear-based device that tracks footsteps from acceleration data and doorways through proximity sensing. From this data, a topological graph of the environment is generated which is applied to localise the user based on a particle filtering approach with preliminary success.

4.2 Eye Tracking based on Periauricular Electrodes

Eye tracking is a widely employed technique for sensing and interaction that typically involves either camera-based [156; 99] or Electrooculography (EOG) [119; 83; 99] methods. EOG detects electric field changes when the eyes move because the eyeballs have a negative charge on the retina and a positive charge on the cornea [289]. While camera-based eye tracking can be precise, it is computationally intensive and requires a significant amount of power [99]. On the other hand, traditional EOG is less computationally demanding and even works when the eyes are closed. However, it is restricted to tracking relative changes in gaze direction, is subject to signal drifts and is relativley invasive as electrodes have to be glued on the face around the eyes. Despite these limitations, EOG has various interesting applications, including hands-free interaction with wearable devices [262; 119], classifying user activities [83], detecting drowsiness from related eye movements [506], quantifying reading activity up to word-level accuracy [81; 253], and providing directed auditory attention in noisy environments [131; 130; 132].

To make EOG more feasible, past work implemented electrodes into smart glasses which improves wearability and more naturally integrates into the life of the user [119; 253]. Prior research has also suggested that eye tracking using electrodes placed inside the ear canal [130; 132; 193; 292; 342; 234] or at the mandible [289] and around the ears [480; 57] is generally feasible. Favre-Felix et al. [132] investigated the use of ear-based EOG and motion sensors around the ear to estimate absolute horizontal eye gaze in multi-talker situations, showing promising results when the head was fixed. However, hardware issues hindered reliable estimations when the head was free. Manabe et al. [292] developed an earphone-based eye gesture input interface using conductive rubber electrodes. In another paper, the same authors [289] proposed a headphone-type gaze detector with electrodes placed at the mandible on one ear to estimate gaze direction. They achieve an estimation error of 4.4° (horizontal) and 8.3° (vertical) in a 5×3 fixation point grid (20° visual angle between fixation points) and lay the foundation for this work. Ear-based eye tracking was also leveraged for sleep detection [374; 342]. Similar to smart glasses, ear-based EOG devices have the potential to be more comfortable, discreet, and portable than traditional EOG. Furthermore, the ear is an ideal location for integrating eye tracking with audio applications for example for directed auditory attention in noisy environments and with hearing aids [130; 132; 193].

The effectiveness in monitoring eye movements using electrodes placed inside the ear canal is well understood, primarily for horizontal eye movements [130; 132; 193; 292]. However, incorporating eye-tracking capabilities into headphones presents several advantages over the in-ear EOG method. Firstly, the proximity of headphones to the eyes enhances the sensitivity to eye movements, leading to potentially more accurate measurements of changes in eye position. Additionally, the larger vertical coverage provided by headphones enables the detection of both horizontal and vertical eye movements, expanding the range of applications and insights that can be gained from this technology. To this extent, this dissertation seeks to expand upon the initial work of Manabe and Fukumoto [289] and evaluate headphone-based eye tracking in depth.

This work thoroughly investigates the hypothesis that EOG-based eye tracking using electrodes placed around the ears in a regular headphone form factor is a reliable and accurate method for studying eye movements. To understand the achievable performance and add context, the research is grounded in comparison to gold-standard EOG. For the evaluation, a specialized headphone device was developed with 28 electrodes positioned strategically around the ears, see Figure 4.2. Using the earEOG headphones, a lab study with 3 participants was conducted to collect data of two tasks.

For the first task, participants were asked to follow one-dimensional moving targets to elicit smooth pursuit data [473]. Smooth pursuits are continuous eye movements that allow the eyes to follow a moving target smoothly without any jerks or abrupt changes in direction [400]. They were chosen for this study because they allow us to identify the best electrodes for detecting eye movements more effectively than saccades, which are rapid and sudden eye movements between fixed points. Smooth pursuits provide a continuous signal that can be more easily correlated with the changes in eye position, enabling a better analysis of the relationship between electrode signals and eye movements. This, in turn, helps us to determine the most effective electrode positions for capturing the nuances of eye movement and to evaluate the overall feasibility of the earEOG method. For horizontal eye movements, the bipolar montage of L4-R4 (difference in electrical signals between the two electrodes) yielded the highest correlation to horizontal gold-standard EOG ($r_{h-L5-R5} = 0.89$). Using electrodes from just one ear, the montage of L5-L10 ($r_{h-L5-L10} = 0.7$) and R5-R10 ($r_{h-R5-R10} = 0.67$) produced the highest correlation. For vertical eye movements, the bipolar montage of L2-L6 and R2-R6 had the highest correlation with vertical gold-standard EOG ($r_{v-L2-L6} = 0.35$, $r_{v-R2-R6} = 0.56$). The position of these electrodes closely resemble the positions of gold-standard EOG.

For the second task, participants were instructed to follow a point that jumped from the center of the screen at 0° to 5° up to 30° in the four cardinal directions at 5° increments. Using the ideal electrode positions identified from the previous analysis it was found that voltage deflections during saccades of earEOG and gold-standard EOG across all angles are mostly highly correlated ($r_{left} = 0.875$, $r_{right} = 0.799$, $r_{up} = 0.019$, $r_{down} = 0.498$). Horizontal saccades exhibit significantly higher correlation to gold-standard EOG than vertical saccades which suggests that earEOG works better for measuring horizontal eye motions. Building upon the relationship between voltage deflections and the underlying gaze angle, a regression model was evaluated to calculate the absolute saccade angle from earEOG. On average, earEOG achieves an absolute angular accuracy of $9.2° \pm 4.5°$. In comparison, a similar model fitted on gold-standard EOG data achieves an absolute angular accuracy of $4.4° \pm 1.1°$. These findings demonstrate that earEOG is a reliable and accurate method for eye tracking, and has the potential to achieve accuracy and reliability close to gold-standard EOG.

In sum, the contributions are: (i) a thorough investigation of EOG-based eye tracking using electrodes placed around the ears in a custom-built headphone form factor, providing earEOG - a novel approach to on-the-go wearable eye tracking; (ii) an evaluation of different electrode positions for earEOG and recommendations for the optimal placement of electrodes, to enable a more effective use of earEOG for eye tracking, thereby enabling a wider range of applications; and (iii) an evaluation of earEOG to predict absolute gaze angles of the four cardinal directions.

4.2.1 Methods

The following sections introduce the participants, apparatus, study procedure, and data evaluation strategy.

4.2.1.1 Participants. The participants in this study were recruited through a sample of convenience. Demographic information was self-reported by participants using questionnaires. Participants without sufficient vision capabilities, as determined by the need to wear glasses, were excluded from the study.

The study protocol was approved by the institutional review board of the

Karlsruhe Institute of Technology (Germany) and followed all relevant ethical regulations in accordance with the Declaration of Helsinki. Before beginning the study, participants were informed of the study protocol, the purpose of data collection, and the specific data that would be collected. Informed consent was obtained from participants through the signing of an agreement form. Participants were not compensated.

4.2.1.2 Apparatus. A 32-channel electric potential data collection setup was implemented using two OpenBCI Cyton + Daisy biosensing boards (one on each ear), with 14 channels around each ear and 2 channels each for vertical and horizontal eye movements gold-standard EOG, see Figure 4.2 (A) to (C). The boards were housed in a 3D-printed headphone enclosure. The electrodes around each ear are metal Ag/AgCl electrodes repurposed from the OpenBCI Biosensing kit. The two boards were manually and precisely time-synchronized using the data captured from the shared electrodes for vertical and horizontal gold-standard EOG. Figure 4.2 (B) shows a schematic representation of the electrode positions on each ear. To enhance conductivity, abrasive gel was applied on each electrode. Before application, the area around the ears was cleaned with isopropanol. Four wet electrodes were placed around the participants' eyes as a reference for capturing vertical and horizontal gold-standard EOG, see Figure 4.2 (C). All data was sampled at 125 Hz.

For data collection, a web-based tool was implemented that provides instructions to the participants and shows on-screen stimuli to elicit specific eye movement patterns. The experiment was conducted using a 23" monitor (1920px $\times$ 1080px) with a viewing distance of 50 cm. Participants were seated centered in front of the screen, with the vertical center of the screen positioned at eye level using a height-adjustable desk. This ensured that the visual stimuli could be presented at the desired size and angle in relation to the participants head and the screen size. The participant's head position was not fixed in space, but the person conducting the study carefully monitored the head position and distance to the screen to intervene if participants moved notably.

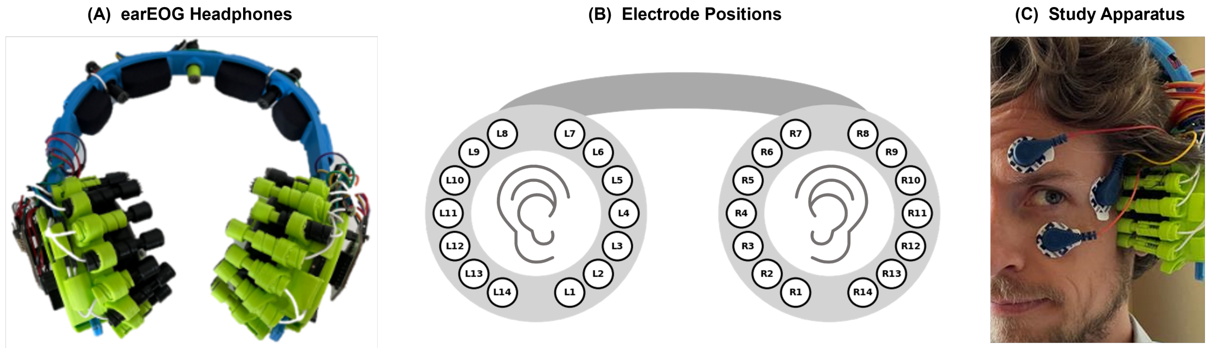


Figure 4.2.: Overview of earEOG headphones showing the electrode positions and study apparatus. (A) The EOG headphones have a standard headphone form factor with 14 gold-plaited electrodes on each ear. (B) Schematic drawing showing preauricular electrodes 1-7 and posterior auricular electrodes 8-14, with L and R indicating the left and right ear respectively; (C) a study participant wearing the EOG headphones and gold-standard Electrooculography simultaneously.

4.2.1.3 Task, Stimuli, and Procedure.

Participants were presented with two different tasks to collect two types of eye movements: smooth pursuits and saccades. Smooth pursuits were added to find the best electrode for vertical and horizontal eye movement tracking, respectively. Saccades were added as they are fundamental for many research studies, and the absolute angle is the most characteristic for understanding eye movements.

Smooth Pursuit Task. Smooth pursuit eye movements are continuous eye movements that follow a moving object, and they exhibit much slower characteristics compared to rapid eye movements such as saccades that shift the gaze from one object or point in space to another. The smooth pursuit task consisted of 1D smooth pursuits angles in both vertical and horizontal directions. Participants were instructed to follow the gaze target that moved within a 5 to 30 degrees visual opening angle from the center for 6 seconds each. The frequency of the gaze target movement was set to 0.33, 0.5, and 1 Hz. The motion of the gaze target was eased by sin(x) on the axis of movement, which means it followed a simple harmonic motion. The gaze target had a diameter of 30 pixels, which is equivalent to 7.8 mm.

The purpose of this task was to find the best electrode for vertical and horizontal eye movement tracking. The eye movement data collected during the smooth pursuit task will be correlated with the eye movements of the gold-standard EOG signals.

Saccade Task. In the saccade task, participants were presented with a series of fixation points to elicit saccade eye motion. Participants started with resting their gaze in the resting position at the screen's center (0° visual angle in the x and y direction). The fixation point then jumped to 5-degree increments in each of the four cardinal directions (left, right, up, down) from 5 to 30 degrees. Each fixation point was presented for 2 seconds and shrank from 30 to 20 pixels (equivalent to 7.8 to 5.2 mm). After completing an angle, the fixation point returned to the center resting point for 2 seconds.

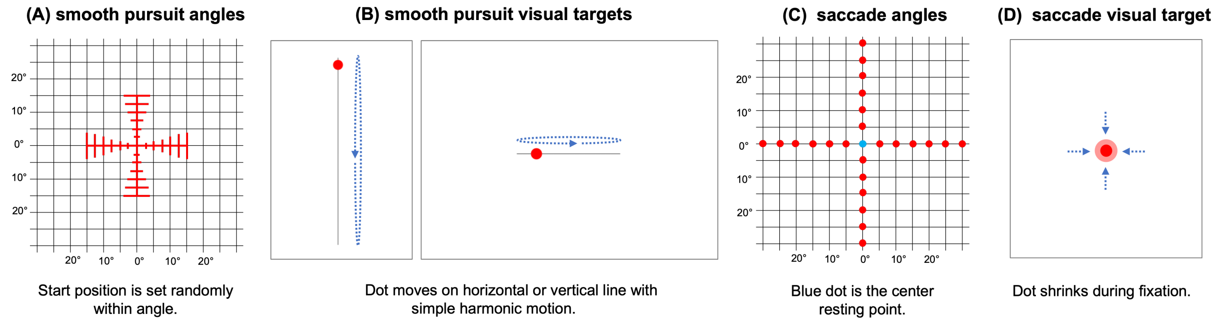


Figure 4.3.: Overview of the different gaze stimuli and tasks presented to study participants. (A) 1D smooth pursuits with gaze target moving on the horizontal and vertical axis with visual angle opening angle of 5° to 30°; (B) visual target moving on a straight line with simple harmonic motion that is shown to participants to elicit smooth pursuit eye motion; (C) fixations points in four cardinal directions with 5° increments and center resting point; (D) visual target point shown to participants that shrinks during fixation.

The purpose of the saccade task was to collect saccade eye movement data that helps to create an understanding of the relationship between earEOG signal deflections, gold-standard EOG, and absolute gaze angle changes.

Data Collection Procedure. After arriving at the lab, signing the consent forms, and asking any possible questions, participants were fitted with the earEOG headphones as well as gold-standard vertical and horizontal EOG electrodes. They were then seated in front of the screen. Participants then followed to on-screen instructions to complete the eye-based tasks for data collection. Both tasks took approximately 10 minutes to complete and were repeated three times, totaling approximately 30 minutes of gaze data per participant. The order of tasks was not counterbalanced.

4.2.1.4 earEOG Analysis. The collected gold-standard EOG and earEOG data per ear are aligned as both were sampled using the same device. To ensure synchronicity between the left and right earEOG data captured by the two OpenBCI boards, the gold-standard EOG data captured by both devices using the same electrodes at the same time were aligned manually by visual agreement of the smooth pursuit data. Pairwise differences between all electrodes were computed for the 14 electrodes on each ear and also between the 14 frontal electrodes between both ears. The difference signals were bandpass filtered between 0.1 and 15 Hz using a 5-th order Butterworth filter to eliminate noise or artifacts. To further smoothen the data, a mean filter was applied with a window size of 10 samples. These preprocessing steps were carried out to ensure that the data was suitable for subsequent analysis. The same filtering was applied to the vertical and horizontal gold-standard EOG.

4.2.1.5 Saccade Angle Prediction. The measurement of voltage deflections through electrooculography (EOG) does not directly provide information on the absolute gaze angle. To overcome this limitation, a regression model was developed that predicts the gaze angle from voltage deflections in earEOG signals. This model was implemented using Python. The 216 saccade samples from 3 participants, 3 sessions, 4 directions, and 6 angles were each labeled manually with a start and end marker. Saccades for which no clear start and end could be identified were excluded (15 in total). From these labels, the voltage deflection of gold-standard and the best earEOG electrodes were computed and assigned to the respective ground truth gaze angle change. This data was used to fit a simple linear regression model from the *sklearn* library. The regression model was evaluated using 5-fold cross-validation. A seperate model was fit to predict vertical and horizontal saccades angles.

4.2.2 Results

In total, 3 male participants were recruited for the study aged between 20 and 30. The following sections present the results and discussion of the evaluation.

4.2.2.1 Comparison of Electrode Positions. In order to find the best electrode positions for measuring eye movements with electric potentials from around the ears, in this dissertation the ground truth gold-standard EOG signal is correlated with pairwise, differential earEOG electrode signals. 1D smooth pursuit data (vertical and horizontal) was used to compute the best electrodes. Smooth pursuit eye movements were chosen because they exhibit continuous electric signal changes compared to rapid saccades, which create a sharp, short spike in the EOG signal. Consequently, smooth pursuits are less prone to artifacts and noise and allow for a continuous sequence to be correlated between the gold-standard and ear-based EOG principle, which improves the reliability of the results.

For horizontal eye movements, electrode combinations for which one electrode is closer to the eye than the other on the horizontal axis were considered (see Figure 4.4 A and B). Similarly, for vertical eye movements, a subset of electrode pairs was analyzed, focusing on those positioned above each other (see Figure 4.4 C). For all recorded smooth pursuits, the full six-second sample was correlated.

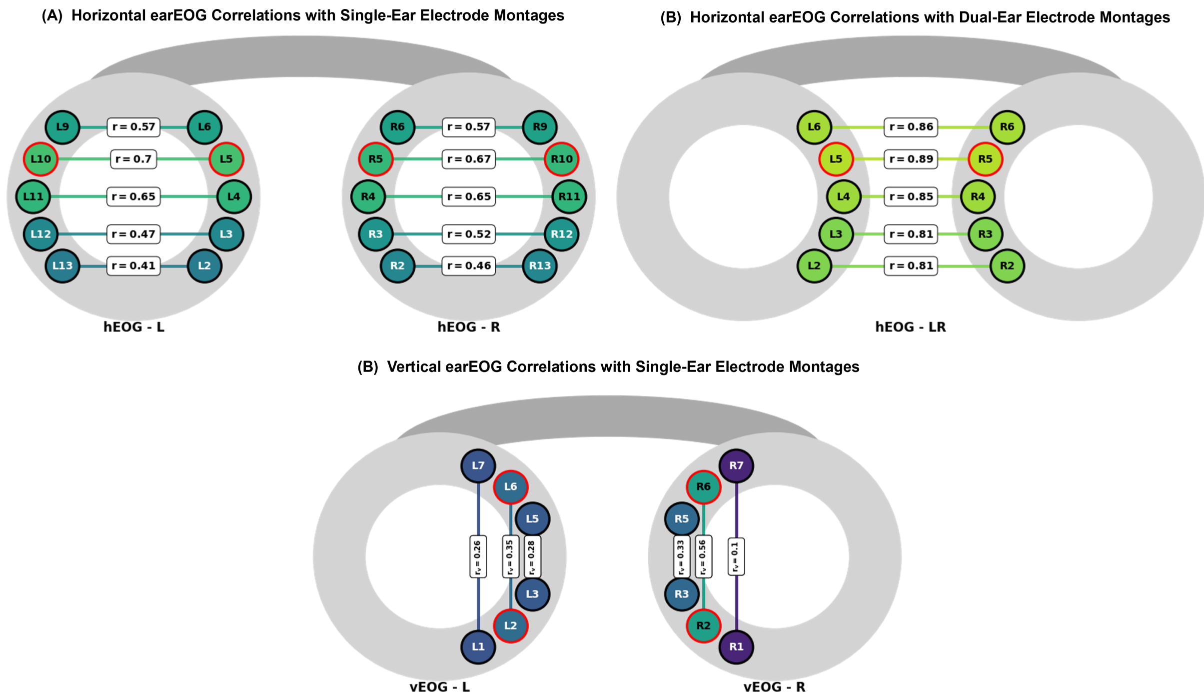


Figure 4.4.: Correlation to gold-standard EOG for different electrode montages. (A, B) Correlations of pairwise electrodes with gold-standard hEOG; (C) Correlations of pairwise electrodes to gold-standard vEOG.

The results are based on all sampled smooth pursuit speeds and angles. To identify the best vertical electrodes, only the vertical smooth pursuit data was used, and for the horizontal electrodes, the horizontal smooth pursuit data was leveraged.

Horizontal earEOG Electrodes. As shown in Figure 4.4 A and B, electrodes at eye level have the highest correlation, and the correlation when combining electrodes from the left and right ear is even stronger than using electrodes from one ear alone for horizontal earEOG with up to $r_{h-L5-L10} = 0.7$ and $r_{h-R5-R10} = 0.67$ vs. $r_{h-L5-R5} = 0.89$. Moving farther away from eye level decreases the correlation. The smallest correlation to horizontal earEOG is achieved by the montage of L2-L13, R2-R13, and L2-R2, which are the farthest away from the eye level.

Vertical earEOG Electrodes. For vertical EOG, the distance to the eyes and sufficient vertical distance covered by the electrodes on the ears are crucial factors. Therefore, the electrodes closest to the ears do not work best; instead, the montage of L2-L6 and R2-R6 yield the strongest correlation (see Figure 4.4 C). Moving farther away from the eyes decreases performance and produces much smaller correlations with L1-L7 and R1-R7 exhibiting the smallest correlation to gold-standard EOG.

Discussion. Horizontal earEOG electrodes showed promising results. Signals were generally stronger when measured closer to the eyes, making the combination of electrodes from both ears advantageous as it captured more of the signal.

Regarding the vertical earEOG electrodes, it was found that the electrodes closest to the eyes were not the most effective. Instead, the L2-L6 and R2-R6 had the highest correlation. This suggests that a sufficient vertical distance covered by the electrodes on the ears is crucial to achieving accurate measurements. The performance of vertical earEOG electrodes diminished as they were placed farther away from the eyes. Overall, the vertical method appeared to be less effective in measuring eye movements compared to the horizontal earEOG.

4.2.2.2 Average Saccade Signal Analysis.

Based on the ideal ear electrodes (horizontal: L5-R5, vertical: R2-R6), the average signal for each saccade direction and amplitude was investigated. Figure 4.5 provides insights into the overall shape and characteristics of the earEOG and gold-standard EOG signals during saccades, enabling a more comprehensive understanding of the eye movement patterns. It is noted, that the signals were shifted to zero based on the average voltage difference of the electrode montage before the saccade was performed to better show the relative change per angle and direction. As to be expected, the gold-standard EOG exhibits stronger characteristics.

Upon examining the average saccade signals for each direction and visual angle, several key features and trends emerge: (i) the EOG signals for saccades of the same direction, irrespective of their amplitude, exhibit a similar waveform shape; this consistency suggests that the EOG system can reliably detect saccades in different directions and that the waveform shape is characteristic of the saccade direction; (ii) the signal magnitude generally increases with increasing saccade amplitude, indicating a strong relationship between saccade amplitude and EOG signal magnitude; (iii) the average saccade signals are reasonably pronounced, allowing for clear identification of the saccades; (iv) vertical saccades exhibit much smaller amplitudes and are less consistent than horizontal earEOG waveforms.

Discussion. The consistency in signal shape, the relationship between signal duration and amplitude, and the overall signal magnitude trends support the notion that the EOG system near the ears can effectively track eye movements across different saccade amplitudes and directions. However, the electric signal properties of vertical eye movements are relatively noisy and less pronounced compared to the horizontal earEOG, which suggests that prediction absolute gaze angles are harder to predict, which is further analyzed in the following sections.

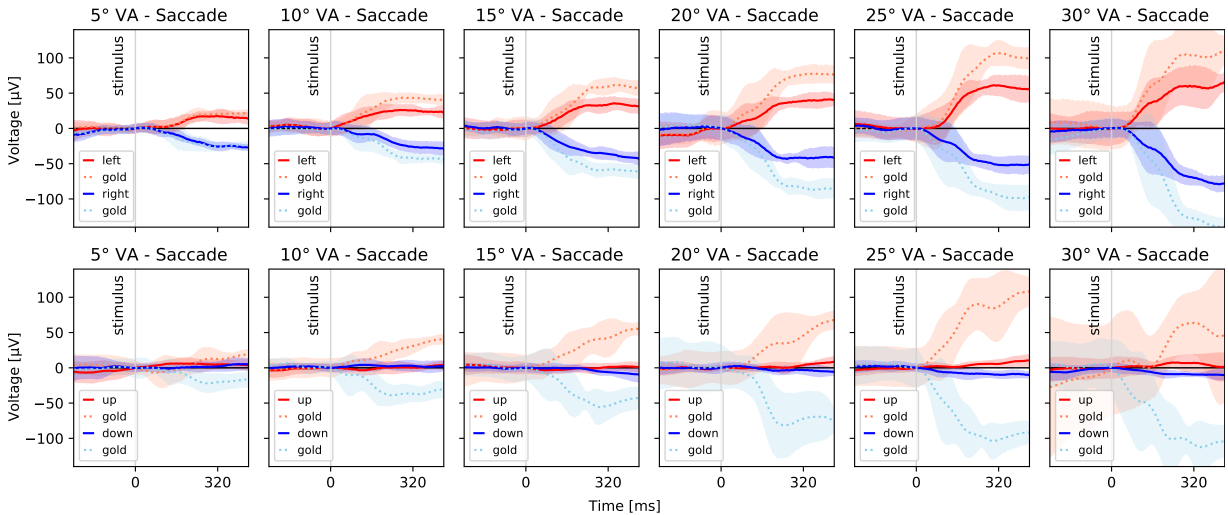


Figure 4.5.: The figure displays the fixation-related EOG signals plotted for different fixation directions (left, right, up, down) and visual angles (5°, 10°, 15°, 20°, 25°, 30°). The red and blue line represent the average earEOG signal. The dotted line in orange and skyblue represent the average gold-standard EOG signal. The shaded area around the lines represents the standard deviation.

4.2.2.3 Saccade Amplitude vs. Voltage Deflection.

The EOG voltage deflections were analyzed for saccades of different angles (5°, 10°, 15°, 20°, 25°, and 30°) in the left, right, up, and down directions. The results are summarized in Figure 4.6 and described below. The correlation between the absolute voltage deflections (ΔV) and the best earEOG positions (h-L5-R5 for horizontal saccades, v-R2-R6 for vertical saccades) and gold-standard wet electrode EOG was also computed.

Horizontal Saccades. For horizontal saccades, the EOG voltage deflections increased with the increase in saccade amplitude. The 5° left saccade showed a deflection of 17.5 µV, while the 30° left saccade resulted in a deflection of 87.7 µV. In the right direction, the deflections were negative, with the 5° right saccade showing a -20.3 µV deflection and the 30° right saccade resulting in a -77.4 µV deflection. This trend indicates that the EOG system was able to differentiate between left and right saccades and capture the increasing deflection with increasing saccade amplitude. For horizontal saccades, the correlations between earEOG and gold-standard EOG deflections was found to be very strong at $r_{left} = 0.875$ and $r_{right} = 0.799$.

Vertical Saccades. For vertical saccades, the EOG average voltage deflections also showed a general increase with increasing saccade amplitude, but the pattern was less consistent compared to horizontal saccades. Upward saccades resulted in positive deflections, with the 5° upward saccade showing a 2.2 µV deflection and the 30° upward saccade resulting in a 7.1 µV deflection. Downward saccades resulted in negative deflections, with the 5° downward saccade showing a -0.8 µV deflection and the 30° downward saccade resulting in a -7.3 µV deflection.

For vertical saccades, the correlations between earEOG and gold-standard EOG deflections were more varied compared to horizontal saccades. For upward saccades, the correlation was only $r_{up} = 0.019$. Downward saccades had a much stronger correlation of $r_{down} = 0.498$.

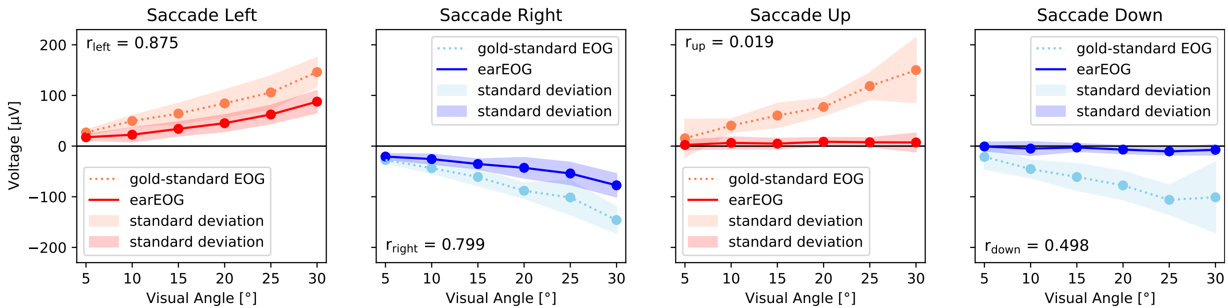


Figure 4.6.: Voltage differences for each visual angle and direction of gaze using earEOG and gold-standard EOG. Scatter plots show the mean voltage differences for each visual angle and direction of gaze using earEOG (red / blue) and gold-standard EOG (coral / lightblue). The filled areas around the lines shows the standard deviation.

Discussion. The results demonstrate that the EOG system applied close to the ears is capable of measuring horizontal and vertical saccades with varying amplitudes. The voltage deflections for horizontal saccades show a more consistent pattern with increasing amplitude compared to vertical saccades. In both horizontal and vertical directions, the EOG system could differentiate between opposite saccade directions, as evidenced by the positive and negative voltage deflections for left and right, as well as up and down saccades, respectively.

The correlation analysis of voltage deflections between the earEOG and gold-standard EOG provides valuable insights into the earEOG system's performance. The high correlations for horizontal saccades indicate that the earEOG system can reliably track horizontal eye movements. On the other hand, the more varied correlations for vertical saccades suggest that there may be room for improvement in the earEOG system's ability to accurately capture vertical eye movements. Especially the upward saccades produced inconsistent results. While the electrodes sufficiently cover vertical distance in the downward direction, the upper electrode is only slightly positioned above eye level. This may contribute to the more inconsistent results.

4.2.2.4 Gaze Angle Prediction. Table 4.1 presents the mean absolute gaze angle errors and standard deviations for the gold-standard and ear-based electrooculography (EOG) methods across different gaze angles (5, 10, 15, 20, 25, and 30 degrees) and directions (left, right, up, and down). The overall error for each method and direction is also provided. The overall errors for the gold-standard method were smaller than those of the ear-based method in every direction: left (3.5° ± 1.3° vs. 7.5° ± 2.6°), right (3.0° ± 1.5° vs. 7.6° ± 2.6°), up (6.1° ± 3.0° vs. 13.5° ± 1.3°), and down (5.2° ± 1.5° vs. 9.6° ± 1.8°).

Discussion. The results indicate that the gold-standard method outperforms the ear-based EOG method in terms of mean absolute gaze angle errors across all gaze angles and directions. The overall error for the gold-standard method was less than half of the overall error for the ear-based EOG method, suggesting that the gold-standard method provides more accurate predictions.

Table 4.1.: Absolute gaze angle prediction results of earEOG and gold-standard EOG. Mean absolute gaze angle errors (± standard deviation) for gold-standard and ear-based electrooculography (EOG) methods across different gaze angles (5, 10, 15, 20, 25, and 30 degrees) and directions (Left, Right, Up, and Down). The overall error for each method and direction is also presented.

Direction	**5°**	**10°**	**15°**	**20°**	**25°**	**30°**	**Overall**
Left (ear)	2.3 ± 0.9	4.8 ± 3.4	5.1 ± 2.4	5.8 ± 2.0	4.4 ± 2.0	6.9 ± 1.3	4.9 ± 0.8
Left (gold)	0.9 ± 0.7	1.9 ± 0.4	3.2 ± 2.1	4.1 ± 2.5	5.1 ± 1.8	5.5 ± 1.1	3.4 ± 0.8
Right (ear)	2.4 ± 1.0	2.5 ± 0.5	3.6 ± 1.6	6.6 ± 2.3	6.9 ± 3.7	5.6 ± 3.7	4.6 ± 1.2
Right (gold)	0.9 ± 0.4	2.3 ± 1.2	3.6 ± 3.5	2.9 ± 2.5	6.0 ± 1.2	3.9 ± 2.5	3.3 ± 1.0
Up (ear)	6.6 ± 2.4	9.0 ± 3.3	10.8 ± 4.8	13.7 ± 7.4	18.9 ± 4.0	24.5 ± 8.9	13.9 ± 2.3
Up (gold)	4.2 ± 4.4	3.7 ± 1.0	6.1 ± 1.1	6.1 ± 0.5	4.5 ± 2.0	10.3 ± 4.1	5.8 ± 1.5
Down (ear)	5.9 ± 3.9	11.8 ± 6.0	12.1 ± 5.5	13.7 ± 2.3	15.4 ± 5.3	21.7 ± 5.9	13.4 ± 1.3
Down (gold)	3.6 ± 2.2	3.4 ± 1.6	4.4 ± 1.7	4.4 ± 2.4	5.4 ± 1.9	10.3 ± 5.5	5.3 ± 1.3
All (ear)	4.3 ± 2.0	7.0 ± 3.6	7.9 ± 3.6	10.0 ± 3.8	11.4 ± 5.9	14.7 ± 8.5	9.2 ± 4.5
All (gold)	2.4 ± 1.5	2.8 ± 0.7	4.3 ± 1.1	4.4 ± 1.1	5.2 ± 0.5	7.5 ± 2.9	4.4 ± 1.1

The increased errors observed in the up and down directions for both methods suggest that these directions may be more challenging for accurately measuring gaze angles. However, the ear-based EOG method showed significantly higher errors in these directions, which could be attributed to factors such as signal interference, noise, or anatomical differences in the placement of the electrodes. Further research is needed to explore these factors and improve the accuracy of the ear-based EOG method in the up and down directions.

4.2.3 Discussion

This part of the dissertation investigated the performance of an earEOG system for tracking eye movements, specifically focusing on smooth pursuits and saccades of different directions and amplitudes. The analysis included the correlations between the earEOG system and gold-standard EOG, the examination of average saccade signals and voltage deflections, and absolute gaze angle prediction. The results of the study provides valuable insights into the potential and limitations of the earEOG system for eye movement tracking and analysis.

Electrode Positions. The presented study identified the optimal electrode pairs for measuring vertical and horizontal eye movements using EOG. It was found that horizontal earEOG works better than vertical earEOG, likely due to the fact that horizontal eye movements are more pronounced and easier to measure, even for gold-standard EOG.

Furthermore, using diagonal electrode configurations for earEOG may be possible. However, upon further investigation it was found that these electrode arrangements are unable to unambiguously differentiate between vertical and horizontal eye movements, as they concurrently capture both dimensions.

The electrode pairs identified by the study can be used in future research and clinical settings conducting similar research with ear-based EOG.

4.2.3.1 Vertical vs. Horizontal earEOG. Overall, the earEOG system demonstrated that tracking horizontal saccades is reliable, with high correlations between the earEOG and gold-standard EOG. The consistency in signal shape, duration, and magnitude trends observed in the average saccade signals and the prediction performance further support the earEOG system's ability to effectively track horizontal eye movements.

However, the results for vertical saccades were more varied, with lower and more inconsistent correlations between the earEOG and gold-standard EOG. This finding suggests that there is room for improvement in the earEOG system's ability to accurately capture vertical eye movements.

4.2.3.2 Comparison to Related Work. Favre-Félix et al. [132] investigated 30° horizontal saccades using in-ear electrodes. They reported a voltage change of approx. 50 μV during saccades. In comparison, earEOG based on the headphone setup achieves 100 μV which suggests that saccades can be measured more reliably using the periauricular positioning of electrodes.

Manabe and Fukumoto [289] used a related headphone setup with gaze targets that were 20° apart. They reported an overall absolute gaze angle estimation error of 4.4° horizontal and 8.3° vertical. Similar to the presented findings, the vertical error exceeded the horizontal error. They calibrate the model per user which increases the performance compared to the general model introduced in this dissertation which is implemented as a one-fits-all approach.

4.2.3.3 Limitations. The current study has several limitations that should be acknowledged. Firstly, the presented work assumes a center resting position for gaze, and therefore does not investigate any relative saccadic movements. As saccades play a crucial role in visual perception, this assumption may limit the generalizability of the presented findings. Secondly, the study does not account for head movements during saccades, assuming instead that the head is fixed in space. However, turning the head is an integral part of gaze, and this oversight may introduce result bias. Lastly, the gaze angle prediction method assumes that saccades have already been detected, and only predicts the angle of gaze. In a real-world system, the isolation of saccades would be a necessary step, and this simplification may lead to an overestimation of the accuracy of the prediction method.

4.2.4 Conclusion

Through the presented research, the potential of using ear-based EOG for measuring eye movements with varying amplitudes and directions was demonstrated. The findings revealed that horizontal earEOG is more effective than vertical earEOG for measuring eye movements. The results also indicate that the gold-standard EOG method outperforms the ear-based EOG method in terms of gaze angle prediction accuracy, particularly in the vertical plane. Nevertheless, the ear-based EOG system shows promising results for measuring eye movements, and future developments may enhance its potential for use in clinical settings and other practical applications such as gaze-based human-computer interaction.

5. Interaction

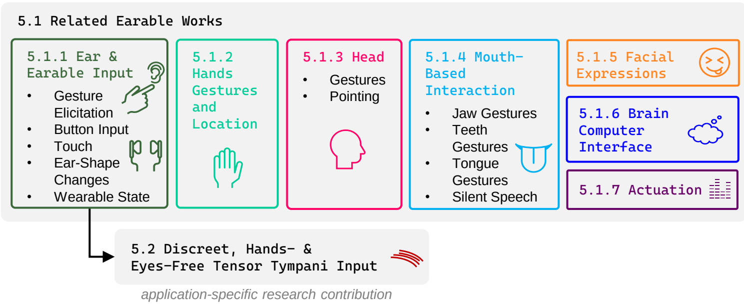


Figure 5.1.: Structure of the “Interaction” section according to different ways of interaction based on sensing on the ears. The section "Tensor Tympani Muscle" is separated out because of the sensing contributions and studies conducted in this dissertation.

Earables present an exciting opportunity for unique and novel interaction techniques given the rich and diverse sensing capabilities available on the earable platform. To date, researchers have explored how input can be provided on the ear or earable itself, as well as how the earable platform can be used to detect other modalities which can be used to provide input, as summarized in subsection 5.1.1. The dissertation expands upon the status-quo through a novel interaction technique based on the voluntary contraction of the tensor tympani muscle, introduced in section 5.2.

5.1 Related Earable Works

The following subsections introduce a series of existing interaction technique for earables, as illustrated in Figure 5.1

5.1.1 Ear and Earable Input

The location of the earable device makes it compelling as an interaction device that can be directly manipulated by the user to provide input. This enables common basic interactions including tap and double tap on and around the ear, however researchers have also investigated how more complicated sliding gestures or manipulation of the ear itself can be utilised for interaction purposes, see appendix A.11.

5.1.1.1 Gesture Elicitation on the Ear. Abstracting from any specific sensing modality, Chen et al. [98] used an elicitation study (N=28) to explore user-defined gestures for ear-based interactions for a number of smart device tasks. The majority of user-defined gestures involved mid-air interactions (57%), with 39% involved touches directly on the ear. Of those touch-based gestures, the most common part of the ear for touch-based gestures was the helix (8.4%), followed by the tragus (and cheek) (6.8%), the lobe (4.7%), the back (4.4%) and finally the center (2%). The remaining touch-based gestures involved multiple ear parts (0.9%) or the location did not matter (11.6%). The study's user-defined gesture set contained a number of touch-based inputs involving single and double taps on different parts of the ear, covering of the ear itself, sliding gestures on the helix, and pinching of the ear lobe. They also discuss the design space of deforming the ear through manipulation, similar to Kikuchi et al. [241]'s *EarTouch*, however these gestures did not feature heavily and the authors speculate that these could cause physical discomfort.

Xu et al. [494] also introduce and explore a rich set of 27 gestures on and around the ear which include single and double tap, as well as simple and complex sliding gestures. They then proceeded to select 8 gestures based on technical properties (signal-to-noise ratio and similarity) of their acoustic-based sensing principle, and user preference based on simplicity, social acceptance and fatigue. Their final gesture set included single and double taps on the cheek, mastoid, and middle ear, as well as two sliding gestures (one on the ear rim and one below the mastoid).

5.1.1.2 Button Input. Pressing a button on an earable device is one of the most basic forms of input. However, the action of pressing inherently requires force to be applied in order to depress the button. In comparison with other parts of the body the ear can be more sensitive to pressure and force which can cause discomfort, especially when applied to the inner canal. Buil and Hollemans [78] found that users (N=16) were split on their preference for the amount of force required to depress a button on an earable, with 85 grams being within the acceptable range of most users.

5.1.1.3 Touch. In contrast, touch-based gestures enable a similar input bandwidth without requiring the same levels of physical force to be applied to the ear. Capacitive sensing has been applied to detect explicit user input using the hands [79; 272; 486]. Buil et al. [79] demonstrated how simple interactions of tap, double tap, and hold could be implemented on the earable platform using capacitive antennas built into the earphones. Lissermann et al. [272] extended this concept to 12 electrodes spanning from the beginning of the ear helix to the ear lobe. They sought to answer how well one can touch their own ear and discovered that users (N=27) are capable of detecting four salient points on the ear arc, with greater precision found at the extrema of the ear arc. Weigel et al. [486] introduced iSkin which demonstrated how capacitive sensing can go beyond traditional capacitive form factors and be achieved using a low-cost, thin, stretchable form factor made up of two layers. Light touch is sensed using the capacitive principle, while firm touch can be detected using the resistive principle when the two layers are in contact due to the pressure exerted. This sensing concept is operationalised in an application called *EarSticker* which consists of a five-element slider located behind the ear, however the sensor itself was not evaluated on the ear.

Due to the importance of real estate in such a small form factor, researchers have explored innovative solutions to detect input by using the built-in components of commodity earphones [290; 494]. Manabe and Fukumoto [290] developed an external adapter which allowed taps on the shell of commodity earphones to be detected through the speaker unit. They tested this technique on ten different pairs of earphones with six users, finding promising results for some earphones, while others did not work due to residual sound. In contrast, *EarBuddy* uses the built-in microphone from earphones to detect a wide range of touches and sliding gestures with an accuracy of $> 90\%$ by utilising deep learning based on mel spectrograms [494]. Input was not limited to just on the ear, but also included interactions around the ear including cheek and mastoid. Similarly, Fan et al. [128] leverage the coupling effect of headphone drivers detected using a peripheral device attached to unmodified headphones. Using this principle, they sense touch and sliding gestures on the earphone enclosure using a cumulative sum algorithm. As a proof-of-concept they found success rates of >99% when audio is off and >97.7% with audio signal with gestures performed 300 times.

In contrast to detecting manipulation of the sensing device, Lee et al. introduced *EarTouch* which uses a camera to detect touches on the ear [259]. The proposed technique was designed to be obfuscated, i.e. hidden from an observer by turning the head, miniaturized due to the small size of the ear, and camouflaged as common actions such as scratching. Their evaluations showed high (approx. 40%) error rates for tapping detection using a "land on" strategy, similar to those found by Lissermann et al. [272], however detection using a "lift off" or "dwell" approach resulted in much lower error rates (approx. 10%).

5.1.1.4 Ear-Shape Changes. Sensing dynamic changes in the shape of the outer ear and ear canal can also be used in the context of interaction. Kikuchi et al. [241] introduced *EarTouch* which uses photo reflective sensors to detect shape changes of the ear caused by physical manipulation of the helix with the hands. A support vector machine classifies four directional gestures of moving the ear helix, which in turn can be used to classify five symbolic gestures (line, check mark, inverted caret, square, and stairs) with an average accuracy of 77.43%.

5.1.1.5 Wearable State. Researchers have also explored how the location of the earables relative to the user can be sensed for both implicit and explicit interactions. Sensing whether the earables are located in the ear can be used for locking the on-ear controls to prevent accidental activation [79], pausing music or answering phone calls [257], or detecting whether the earphones are positioned correctly (i.e. left ear bud in left ear, right ear bud in right ear) [302].

The wearable state of the earables can be detected using capacitive sensing [79], proximity sensors [302], ultrasound frequency sweeping [257], and the coupling effect through the headphone drivers [128]. *HeadFi*, introduced by Fan et al. [128], provide an extensive study on detecting the wearable state of the earphones using 54 pairs of earphones grouped into five types. By leveraging the coupling effect through the headphone drivers they achieved success rates per group between 97.9% and 99.8%.

Beyond detecting just whether the user is wearing the earables or not, *EarphoneTrack* uses acoustic motion tracking to find out where the earphones are located in 3D space with millimeter level accuracy [85]. The system leverages an inaudible single frequency acoustic signal and can be used with commercial earphones. They propose self-interference and frequency offset techniques to allow for the tracking of both wired and wireless earphones respectively. This approach creates an interesting and untapped design space where the earables themselves can be used for spatial input.

5.1.2 Hand Gestures and Location

Besides direct manipulation of the earable or the ear for input, hand movements can also sensed and leveraged for input. Using an elicitation study (N=28), and abstracting from any specific sensing modality, Chen et al. [98] found that the majority of user-defined gestures for ear-based interactions involved mid-air hand interactions (57%) around the ear. Other researchers have operationalised sensing mechanisms which can detect motion of the hands using earables, see appendix A.12.

Metzger et al. [312] introduced *FreeDigiter* which allows rapid, contact-free entry of digits based on finger gestures. To enter a digit, the user spreads their fingers to show the desired number and then slide it over a proximity sensor embedded in the earbud to encode the digit. The earbud detects the digit by the reflections of the infrared light emitted by the proximity sensor which can be used for selecting a numbered item from (e.g.) a list of tracks or to manage phone calls. *SonicASL* by Jin et al. [215] uses deep learning techniques to classify reflections of a sonic wave into 42 different sign language words. This technique enables mid-air input in front of the earable that can be used to enable communication with hard-of-hearing individuals. Tamaki et al. [443] presented *Brainy Hand* which embeds a mini-projector and a color camera inside an earbud capable of detecting hand gestures. The projector gives the user feedback where input can be performed (camera field of view), and in another configuration the projector displays words and images on the user's palm for richer feedback.

Beyond detecting mid-air hand movement for explicit user input, hand movements have also been detected in more subtle and novel ways. Yan et al. [497] introduced *PrivateTalk*, a subtle interaction technique to activate voice input by partially covering the mouth with the hand from one side. This action causes differences of the audio signals arriving at the left and right ear which can be used to signify that the user intends to interact with the voice assistant when the hand covers the face. This removes the necessity to use wake-up words or pressing a button while increasing the privacy of the user by reducing the spread of voice and concealing lip movements.

5.1.3 Head Gestures and Pointing

Head movement provides a hands-free input mechanism when the hands are busy or unavailable. The head has been used for input by either semantically mapping pre-defined gestures to system commands (e.g., [496]), or by using the direction as a pointing device to select targets spatially (e.g., [412]). Earables are perfectly situated to detect head movements, and researchers have explored how sensing on the earable can be used to operationalise both of these interaction paradigms.

5.1.3.1 Head Gestures. In addition to being hands-free, head gestures, such as nodding or shaking, are compelling because they can also be invoked in an eyes-free manner due to proprioception, and do not require visual feedback. The inertial sensors within an earable can be used to detect different head gestures intended for interaction. Gashi et al. [153] combines a hierarchical classification with transfer learning to detect typical activities including head shaking and nodding from accelerometer and gyroscope data. They achieved an F1-score of 88.24%. Similarly, Laporte et al. [256] detected nodding and shaking (and three other activities) with an end-to-end deep learning approach with an F1-score of 82%. Rather than using inertial sensors, Ando et al. [16] leveraged the fact that the ear canal changes shape when the sternocleidomastoid muscle is used to move the head. They used in-ear pressure sensing to recognize six head gestures (rotate left/right, rotate up/down, tilt left/right) with a recognition accuracy of 87.6% (which also included five facial gestures). While these papers show a promising future direction for earables interaction, they also highlight a common challenge with semantic gestures – distinguishing movement intended for interaction versus natural head movement.

5.1.3.2 Head Pointing. Pointing is a fundamental interaction principle that is at the core of graphical user interfaces. An inertial measurement unit that can track yaw, pitch, and roll can be used as a pointing device with three degrees of freedom. However, a magnetometer is required to measure the absolute yaw position which has proved problematic for earables due to electromagnetic noise in such a small form factor [136]. Odoemelem et al. [350] used head motion to control a robot arm, however due to the lack of yaw information of the head they map the roll of the head to the yaw of the robot. Two research threads have emerged due to the difficulties with the magnetometer: the first seeks to understand the source of the error and overcome it using calibration techniques, and the second thread seeks alternate ways of detecting the yaw position of the head.

Ferlini et al. [133] aimed to compensate for the lack of magnetometer by utilizing an additional gyroscope. Instead of detecting a specific yaw angle of the head, they classified from 30 degrees to 90 degrees in 15 degree increments. They demonstrate errors of between 5 - 15 degrees between the ground truth and proposed approach, which increases when the user is chewing or speaking. The prediction error also increases when the user maintains their position due to sensor drift. More recently, Ferlini et al. [136] demonstrated how an automatic magnetometer calibration method can overcome the electromagnetic interference resulting in an error of less than 5 degrees over a wide range of yaw angles. This approach may be key to future opportunities for earable-based head pointing.

As a result of difficulties with head pointing on the earable platform, other work has looked at innovative ways of leveraging the earables relationship to other devices to infer head direction. Hashem et al. [181] introduced the *Look&Lock* system which determines which device a user is looking at by using the Bluetooth received signal strength (RSS) on a set of earables.In a single subject study, the system was capable of 100% accuracy when detecting objects on the walls or in the corner, spaced by 15 degrees, at short distances of less than 3m. Closely related, Pfreundtner et al. [373] used the same principle with audio signals instead. They used four microphones (two on each ear) to estimate the direction of a sound source in relation to the head with an accuracy of 14 degrees on the horizontal and 5 degree on the vertical plane. Whereas the previous two papers used fixed devices in the environment, Gamper et al. [149] tracked the head orientation between multiple earable wearers by taking the speaker's voice as sound emission source and estimating the relative head angle to the listener using binaural microphones with an accuracy of around 10 degrees. This set of approaches may provide interesting opportunities for cross-device interactions, however they do not provide the accuracy required for pointing selection on a single device.

5.1.4 Mouth-Based Interaction

Similar to head movements, mouth-based movements afford hands- and eyes-free opportunities for input. The mouth provides a surprisingly rich input space for interaction [97] involving movement of the jaw, clicking of the teeth, and positioning of the tongue. appendix A.13 provides details on the different systems that have been used for explicit input using the jaw, teeth, tongue.

5.1.4.1 Jaw Gestures. Physiologically, jaw movements can be detected on the ear due to changes in the shape of the ear canal, however the magnitude of these changes can strongly vary between users [162; 378]. More specifically, when the jaw moves the shape of the ear canal changes depending on the position of the mandibular condyle [16]. Researchers have explored how the resultant ear canal deformations can be sensed using in-ear pressure sensing [16], proximity sensors [46], or piezoelectric bending sensors [88].

Ando et al. [16] explored how sensing in-ear pressure can be used to detect a wide variety of facial and head gestures due to the ear canal changing shape. They showed that gestures involving sliding the jaw left and right, as well as basic open and closing of the mouth, can be detected with >88.2% accuracy against other head movements (which also cause ear canal deformation). They also explored how different levels of mouth open can be detected, with four levels (closed, slight, open, wide) showing a minimum accuracy of 79.2%.

In contrast, Bedri et al. [46] used three orthogonal proximity sensors to detect the change in shape of the ear canal, and Carioli et al. [88] used bend sensors on a custom fitted ear piece. Proximity sensors were used in the context of silent speech recognition [407] and eating detection [48], but in theory both sensing modalities could be used for interaction.

5.1.4.2 Teeth Gestures. Movement of the jaw can also involve clicking the teeth together which results in vibrations and an audible sound. Tooth-based interaction has been used as an activation gesture [435] as well as for navigating menus [22; 389], answering phone calls [435], and typing on a keyboard in both assistive and non-assistive use cases (e.g., while working out) [22]. Such smaller interactions reduce the effort required to perform short input tasks [22] and maintain the privacy of the user [389], however users reported issues of jaw muscle fatigue when using teeth for typing [22].

Tooth-based interaction is commonly sensed using audio-based approaches to classify the distinctive sounds and vibrations from teeth clicking, similar to eating detection in subsection 4.1.4. Ashbrook et al. [22] reported a recognition accuracy of 96% for five different tooth pairs using bone conduction microphones, showing that the location of the click can be determined. Prakash et al. [389] expanded the input space for tooth-based interaction by exploring sliding gestures and found that these can be distinguished from taps with >90% accuracy (using six gestures). Sun et al. [435] used a fusion of audio-based and inertial sensors to detect 13 gestures, consisting of hold gestures combined with single, double, and triple taps with 90.9% accuracy in a lab environment.

5.1.4.3 Tongue Gestures.

Interestingly, movement of the tongue can also be sensed on the ear as a result of the deformation of the ear canal [86; 287; 444]. This principle can be used for private input techniques when in a public setting [344; 444], and provides an accessible means of input for users who have a speech impairment [287; 344; 86] or physical disability [344; 86]. Interaction can be based upon pointing with the tongue in pre-defined directions within the mouth, or by detecting whether the tongue is protruding or retracted in the mouth [287; 444; 344; 86].

Earable form factors have shown how detecting tongue movements is possible. Maag et al. [287] use in-ear pressure sensing to detect ear canal deformation of three tongue gestures (left, right, and front) and two interfering movements (removing device and moving head) with comparable accuracy but significantly lower power requirements than other audio-based approaches. Taniguchi et al. [444] use a miniaturised optical sensor to detect when the tongue pushes the roof of the mouth in a small (N=5) feasibility study which shows promising results.

Other work, has explored how sensing more advanced tongue movement is possible at the ear. Nguyen et al. [344] built a more complex setup from multiple sensors including EEG, EMG, and skin surface deformation and demonstrated the feasibility of detecting ten different touch points using the tongue that can be used to provide input. Their setup achieved >85% accuracy for eight out of ten locations including four on the tongue-side (lingual) of the teeth, and six on the cheek-side (buccal). Participants found the cheek-side movements easier than the tongue-side, however many found the technique difficult to use. Finally, Cao et al. [86] use a smartphone held to the ear to sense ear canal deformation by sensing acoustic reflections measured using the microphone. Both of these are promising avenues, however it is unclear how well they can be translated to a practical earable form factor.

5.1.4.4 Silent Speech Input.

Silent speech recognition is a unique use case of detecting both jaw and mouth movement in synergy on an ear-based device. Silent speech offers the user a privacy-preserving, socially acceptable interaction technique [237; 96] which can be used in noisy environments [407; 96] and by users with medical conditions [407].

Ear canal deformation can be used to detect silent speech, similar to jaw and tongue movements. Khanna et al. [237] developed *JawSense*, an accelerometer-based approach that uses ear canal deformations to classify 9 phonemes with 92% accuracy (N=6). Sahni et al. [407] combined an earable equipped with proximity sensors and a magnet attached to the tongue, sensed by a magnetometer, to detect 11 sentences at 90.5% accuracy (N=6). Chen et al. [96] opted to sense cheek deformation instead of ear canal deformation. They used a camera attached on the ear of a user to predict eight words with 84.7% accuracy (N=6). These early works demonstrate how silent speech recognition is possible on the ear, however there are still open research questions and challenges with regards to using and deploying this technology in-the-wild.

5.1.5 Facial Expressions

Facial expressions can either be voluntary or involuntary. Voluntary facial expressions are a form of non-verbal communication and can be used for explicit input for mobile and wearable devices [304; 11; 411; 96]. Involuntary facial expressions can be used in affective computing to capture the underlying feelings, mood, or emotions of a user and can be used as implicit context Verma et al. [472]. Additional insights into emotion detection with earables can be found in subsubsection 3.1.4.1). appendix A.14 details a complete list of the different sensing principles and facial expression gestures explored with earables.

The muscles which control facial expressions (i.e., sternocleidomastoid, masseter, and temporalis muscles – see Figure 2.4) all run close to the ear. Preliminary work has shown how in-ear EMG can be used to sense the muscle contractions in the ear when someone performs a facial expression [411]. Facial expressions can also be sensed indirectly using the deformation of the ear canal. Matthies et al. [304] explored the use of electrical field sensing which, based on a single user, performed similar to EMG in a lab environment, but when studied in a mobile context with 3 users, the electrical field sensing approach was more robust resulting in better performance. They also found that placing non-insulated electrodes in a vertical arrangement produces better results for detecting the in-ear deformation caused by facial expressions

Amesaka et al. [11] sensed the deformation of the ear canal using reflected sound, by measuring the ear canal transfer function with a microphone in the earable device. They show that it is possible to achieve a recognition accuracy of 62.5% when classifying 21 gestures, increasing to 90.0% with a smaller subset of 6. This approach of exploring a larger number of facial expression gestures (20+) before deciding on a smaller subset to increase the recognition accuracy is common in the literature [304; 11; 264].

Similarly, inertial data was applied to extract characteristic motions during facial expressions [472; 264; 153]. Verma et al. [472] reported very high accuracy (89.9%) for up to 32 facial expressions when training a user-specific classifier, however this performance decreased to 42.1% for a user-independent model. Ear-mounted miniature cameras can also be used reconstruct the outline of the cheeks with a mean square difference of 0.77mm and 0.74mm for both earphones and headphones respectively, with little degradation in performance when the user wears a mask (0.717mm) or glasses (0.824mm) [96]. Despite these promising results, the applications of camera-based approaches may be limited as they require significant power and raise potential privacy concerns of passers-by.

5.1.6 Brain Computer Interfaces

Brain-computer interfaces (BCI) offer input based on brain activity alone. Commonly, users have to execute a specific task or follow a stimulus that triggers an expected response of the brain which can be quantified by electroencephalography (EEG), which was evaluated broadly with earables (see subsubsection 3.1.2.1). In the context of BCI, the steady-state response of the brain plays a critical role as it creates a response in the EEG trace that matches the frequency of an auditory or visual stimulus attended by the user. By offering multiple stimuli at the same time and at different frequencies, a user can select a desired option by attending the specific stimulus [279; 58]. Accuracies of 79.9% (six visual stimuli, [5]), and 87.92% (four visual stimuli, [482]) were achieved with in-ear electrodes. Based on the principle, text-spelling was possible at 2.4 characters per minute [348]. Brain activity also depends on the task a user is performing. Up to 90% binary selection accuracy was achieved when selecting the two ideal tasks among a list of 5 activities (breath, imagine song, listen to tone, imagine a face, imagine a cube rotating shown on screen) [311].

5.1.7 Actuation

Earables commonly use auditory output to share feedback with the user, or rely on visual feedback from an external device. Visualisations on the earable can not provide information to the wearer directly during use due to the location of the ear, however visual cues, such as colour-changing LEDs, have been used to provide feedback to other users [100]. Beyond these common output modalities, researchers have explored how ear-based devices can provide alternative output mechanisms based on haptics, thermal cues, and mechanical deformation.

5.1.7.1 Haptics. The ear is one of the best body locations for perceiving vibrating stimuli based on research which has shown that vibrations on the ear had the highest perceivability compared with other body parts including the hand, foot, and neck [120]. However, other work suggests that users could better perceive vibration stimuli at the ear because they could hear them [487]. To make full use of the ear as a tactile display, Lee et al. [263] performed a thorough investigation of vibration stimuli at the lobe, concha, and superior crus (under the helix) on each ear. Across all locations, users perceived 15 Hz to be the clearest and most unobtrusive vibration frequency. Additionally, sequential stimuli were easier to perceive and distinguish than simultaneous stimuli, while 25 spatio-temporal dual-ear patterns yielded 58.2% recognition accuracy with 4.8 seconds average response time. Vibrotactile feedback at the ear may also give hard-of-hearing users the ability to understand the sound of their surroundings better. For example, they could identify simple words and environmental sounds from vibrations at different intensities depending on the frequency of the incoming sound [487].

5.1.7.2 Thermal Cues. Hot and cold sensations can be created around the ear using Peltier elements [8; 338; 339]. Akiyama et al. [8] initially presented how thermal sensation changes around the ear may enhance the excitement and comfort of users listening to music. An in-depth evaluation of the thermal cues revealed that four periauricular thermal cues could be distinguished reliably by users at >99% recognition accuracy, while 5 parallel cues significantly reduced recognition performance to 86% [339]. Superior auricular cues were perceived less accurately due to the user's hair, while the posterior auricular area was most sensitive. Spatio-temporal combinations of cues on one or two ears resulted in 14 patterns with 85.3% recognition accuracy and 2.3 seconds average response time.

5.1.7.3 Ear Deformation. Inspired by the movement of the ears during communication of animals, Huang et al. [204] introduced a system that applies mechanical actuators to deform the ear with a view to extending the body language of disabled users. They propose 22 static and dynamic auricular postures which they link to different emotions based on an online survey. Closely related, Shirota et al. [423] applied linear actuators behind the ear to change the opening angle of the ear which could successfully manipulate the perceived direction of sound.

5.2 Discreet, Hands- & Eyes-Free Tensor Tympani Input

Earphones are widely adopted by consumers because they provide private audio channels to listen to music, podcasts, or audiobooks, to watch the latest movies and TV series whilst commuting, or to make hands-free phone calls. As shown in the previous sections, earables present new opportunities for interaction with mobile devices - where interactions may occur when the user is pre-occupied, or in public spaces. Input techniques using "subtle" or "motionless" input gestures are desirable in mobile contexts because they take into consideration the social context of mobile device usage [108].

Input techniques with little to no movement avoid the inconvenience of techniques requiring large physical effort (e.g. hand gestures) [191] and are more socially acceptable to spectators [398]. The latter benefit also has the advantage from a user's perspective of maintaining a level of privacy over the interaction to avoid unwanted attention [383]. Microgestures and hands-free input approaches enable users to interact with their device without disrupting other tasks they may be performing, for example manual tasks such as writing a letter or driving a car [492]. Earables can be used to detect input by tapping the earphone itself [80; 291], or as a sensing platform to detect more advanced gestures performed directly on [242], around [495; 271], or in front of [313] the ear (see previous section 5.1 for details). The earable platform also presents unique opportunities for interaction, such as the use of in-ear barometry to detect head gestures [17]. However, the use of such interactions, in addition to others such as opening and closing the mouth, raises social acceptability issues [398] and there is large scope for the exploration of more discreet methods of input using the earable medium.

This dissertation introduces *EarRumble*, an interaction technique which uses contraction of the tensor tympani muscle inside the middle ear, see Figure 5.2 (a). The tensor tympani muscle is the second smallest muscle found in the human body, and is used for dampening loud noises [218]. Interestingly, a subset of the population has voluntary control over the contraction of the tensor tympani muscle which causes in-ear vibrations when contracted [384; 431]. As the tensor tympani contracts it tightens the eardrum and the volume encapsulated in the ear canal rises, resulting in a reduction in pressure. EarRumble measures pressure changes within the sealed ear canal to unobtrusively detect contraction of the tensor tympani [197], see Figure 5.2 (b).

The dissertation explores how this phenomenon, also known as ear rumbling, can be exploited for interaction with mobile devices using simple, discreet gestures. However, not everyone has voluntary control of the tensor tympani. To inform how many people could hope to use ear rumbling for interaction, and to demonstrate that rumbling is a viable opportunity to pursue, this dissertation first investigates the prevalence of rumbling through an online questionnaire (N=198) which revealed that 43.2% of respondents had the ability to voluntarily contract the tensor tympani. As contraction of the muscle is unlikely to occur voluntarily on a regular basis, it is unclear what level of control users have over the muscle for the purposes of interaction. To address this, this work provides initial insights into the complexity of the interaction design space through analysis of different gestures (e.g. sequential rumbles, holding the rumble).

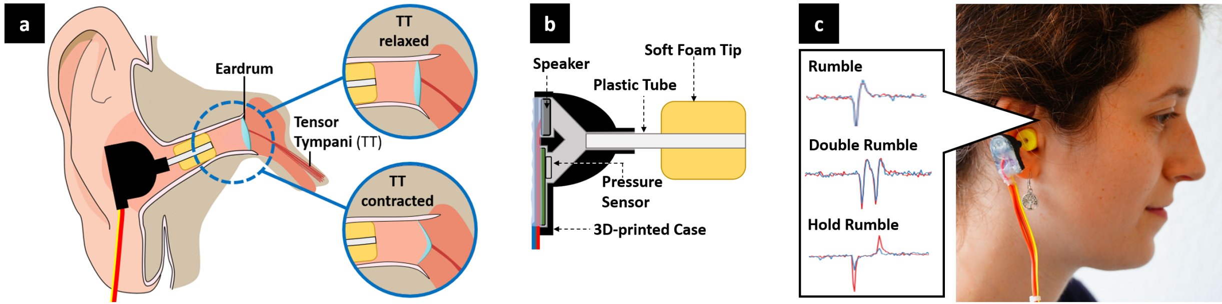


Figure 5.2.: The dissertation presents EarRumble, a technique that uses “ear rumbling” for interaction. (a) The tensor tympani muscle can be contracted voluntarily which displaces the eardrum and induces a pressure change within the sealed ear canal; (b) Custom-built earables detect ear rumbling using an in-ear pressure sensor; (c) Eyes- and hands-free discreet input can be provided by performing different rumbling gestures by voluntarily contracting the tensor tympani muscle.

The dissertation also gathers feedback on user perception of how easy and comfortable different rumble gestures are to perform on demand to determine their viability. In this dissertation, data from participants with the ability to ear rumble is collected (N=16), and the characteristics of three ear rumble gestures is analysed, see Figure 5.2 (c). Using the data collected, a recognition pipeline is implemented which detects ear rumble gestures from everyday activities that may also induce pressure changes in the ear canal, with up to 95% accuracy with real-time performance. Finally, EarRumble is explored to be used for interaction in two exemplar applications (receiving phone calls and an audio player) using three manual, dual task application scenarios that one might face when using earables in everyday life. Users (N=8) provided positive feedback, describing how the EarRumble technique felt "magical" and "telepathic", and highlighting how EarRumble is a low-effort, hands- and eyes-free input technique for earables.

In sum, this part of the dissertation provides the following contributions: (i) EarRumble, a hands- and eyes-free, discreet input technique based on voluntary control of the tensor tympani muscle found in the middle-ear, and sensed through in-ear barometry; (ii) An indication of how prevalent ear rumbling is, how easy it is to perform on demand, and how comfortable it is through an online questionnaire; (iii) Data-driven insights into how well users can contract the tensor tympani muscle in the context of interaction; and (iv) Insights into how EarRumble can be used for interaction, grounded in real-world applications involving dual task scenarios. The presented work was published at CHI 2021 [453].

5.2.1 Background and Related Work

First, subtle and discreet interaction in HCI research is introduced. Also, the medical background of the tensor tympani muscle is provided. For a summary of related techniques for interacting with earables, readers should refer to section 5.1. None of the work that was identified regarding interactions on, around, or by the ear investigated voluntary control of the tensor tympani muscle as an active input mechanism which, therefore, presents an entirly new technique introduced to HCI.

5.2.1.1 Subtle and Discreet Interactions. This dissertation introduces a subtle and discreet interaction technique that requires low effort and can be hidden from others – two areas that were recently highlighted as part of a systematic investigation into subtle interaction in the HCI literature [383]. Motivations for *doing less* in interaction include (a) to make interaction smaller and more comfortable [275] so that they do not cause physical discomfort [117], (b) being always available [275] and/or (c) "*to execute a secondary task, for example controlling mobile applications, without interrupting the manual primary task, for instance, driving a car*" [492]. Costanza et al. promote the term *intimate* interfaces, meaning subtle, discreet and unobtrusive control of mobile devices [108]. Systems that enable subtle interaction, but involve technology that itself is not subtle include *Gunslinger* (two 3D cameras for bare-hand gestures) [275], the *Magic Ring* (finger-worn wearable with accelerometer to detect small finger gestures) [216] and *Bitey* (teeth clicking through a head-mounted bone microphone) [23]. In contrast, the approach in this dissertations integrates pressure sensors with commodity earphones, and the act of contracting the tensor tympani does not necessarily require an externally noticeable user gesture or facial expression. Other discreet interaction techniques of note are *Itchy Nose* by Lee et al. which employed EOG sensors embedded in the frame of smart glasses to detect small finger gestures performed on the nose [260], and Gallego Cascón et al.'s *ChewIt* – an intraoral device that resembles an edible object and allows hands-free input-operations [148].

Another common motivation for subtlety is to enable socially acceptable interaction, meaning to not "*disrupt [...] others in the vicinity, or others in the group*" [396]. Users may desire privacy, e.g. to protect private texts, passwords, or PIN entry. Taking this to extremes can mean completely hiding the fact that interaction happens at all. There are also application-specific motivations for subtle interaction: many researchers have approached discreet interaction in the context of different modalities, ranging from micro-gestures with the hand [145], fingers [93], gaze [114; 231], and oral interfaces [466; 148]. The use of the ear is underexplored, and EarRumble enables interaction in which users can *do less* and are not impaired in other actions, as well as provide *hidden* interaction that is undetectable by, and non-disruptive to, others.

5.2.1.2 Tensor Tympani Muscle.

The tensor tympani muscle sits in the human middle ear and actuates the tympanic membrane during the middle ear reflex [431; 327]. The tensor's subconscious contraction accompanies vocalization and swallowing [245], or expecting a startling sound [122], and the ability to voluntarily contract the tensor tympani muscle has been discussed for over one hundred years [385]. Due to the vibrations induced by the tensing muscle, a rumbling sound can be heard [385], and those who can voluntarily contract the muscle often describe it as flexing, activating, or moving a muscle inside their ear[1]. The sound might be imitated by firmly tensing one's fist and pressing it on the ear to create a comparable dull rumbling sound. Little information is available about the prevalence of the ability to consciously control the muscle [18]. In this dissertation, an online questionnaire is conducted to provide insights into the prevalence of voluntary contraction, and to explore the applicability and constraints of an input technique that relies on voluntary tensor tympani control.

The medical field documented multiple principles to sense tensor tympani contraction [327]. Electromyography is the most invasive method as it requires surgery for placing electrodes [408]. Otologists commonly measure the acoustic impedance of the ear which can be used to detect contraction [327; 326]. However, the technique requires playing an 800 Hz probing tone. It is also possible to use a camera otoscope to detect contraction. This is used in subsection 5.2.4 for validation that participants are contracting the tensor tympani by visually observing the eardrum displacement as a result of tensor tympani contraction. However, camera-based technology is not suitable because of power consumption, computational complexity, occlusion (e.g. ear wax), and the need to focus the camera on the eardrum. Alternatively, in-ear barometry can be used to measure the displacement of the eardrum during contraction through pressure changes in the sealed ear canal [197]. EarRumble uses this approach because it does not require playing a constant tone, is cheap to realise with off-the-shelf components, and can be incorporated into a pair of earphones. Previous work has also demonstrated the utility of in-ear barometry for interaction through head gesture and facial expression detection [17].

[1] Ear Rumblers Assemble subreddit: `www.reddit.com/r/earrumblersassemble/`

5.2.2 EarRumble

The following sections introduce *EarRumble*, an interaction technique based on the voluntary contraction of the tensor tympani muscle found in the human middle ear - a phenomenon also known as ear rumbling. The contraction of the tensor tympani is detected using in-ear barometry, i.e. measurement of pressure changes within the ear canal, using a custom-built earable consisting of commercial off-the-shelf (COTS) components and a custom 3D-printed enclosure, see Figure 5.3. A thresholding detection algorithm and feature-based machine learning classifier are applied to recognise three basic ear rumbling gestures from the raw pressure signals, see Figure 5.4.

5.2.2.1 Concept. Figure 5.2 (a) illustrates the underlying principle of the EarRumble interaction technique. Upon contraction of the tensor tympani muscle, the eardrum displaces inward. As the soft foam earcaps worn by the user seal the ear canal, the volume increases while the encapsulated air remains constant which results in a pressure drop. After the relaxation of the muscle, the eardrum returns to its original position. This leads to a pressure wave that is pushed outwards of the ear canal to produce a positive pressure peak, as shown in the *Rumble* signal in Figure 5.2 (c).

As the sealing of the ear canal is not perfectly air-tight, the pressure equalizes over time. Therefore, holding the tensor tympani contracted does not yield a constant low pressure reading, however releasing the muscle still produces a sufficiently pronounced pressure peak in the opposite direction, as shown in the *Hold Rumble* signal in Figure 5.2(c). EarRumble utilises the changes of pressure to derive contraction events of the muscle which can be measured with a standard, off-the-shelf pressure sensor. The presented technique assumes that the tensor tympani can be in one of two states – relaxed or contracted – and the exploration focuses on the use of three basic ear rumble gestures based on insights from sections 5.2.3 and 5.2.4:

- Single rumble – a quick contraction of the tensor tympani
- Double rumble – two contractions of the tensor in quick succession
- Hold rumble – contraction of the tensor tympani for approx. one second

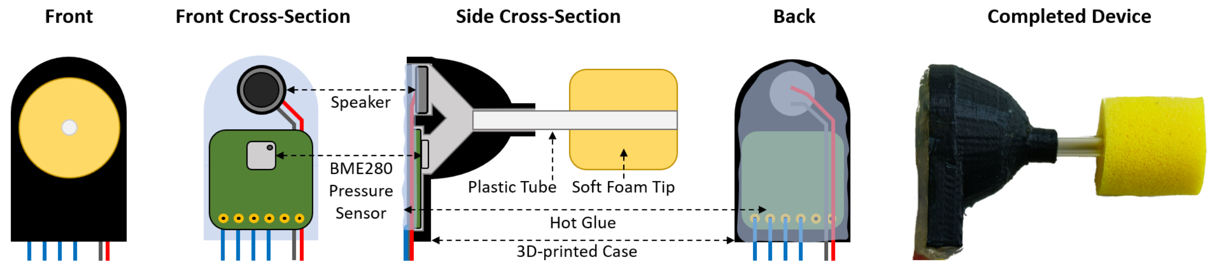


Figure 5.3.: The custom-built device used to realise the EarRumble technique. Each earbud contains a speaker, which provides the usual audio capabilities of a pair of headphones, and a BME280 pressure sensor which is used to detect the ear rumble gestures by measuring the changes of pressure inside the ear canal. The soft foam tip and hot glue provides an air-tight seal within the ear canal. A custom 3D-printed enclosure houses the components and provides two separate channels for the speaker and pressure sensor.

5.2.2.2 Hardware. Figure 5.3 illustrates the assembly of a custom-built device. To realise in-ear barometry, Bosch BME280 pressure sensor (2.5 x 2.5 x 0.93 mm) sampling at 32 Hz is used. The speaker was removed from a pair of commercially available earbuds (Sony MDR-EX110LP) and a custom earplug case was 3D-printed that encapsulates the different components in a single device. The case splits the enclosed air into two channels, one directed towards the speaker and the other towards the pressure sensor. This was done to minimise the volume of air enclosed within the ear canal, and maximise the change in pressure. To ensure tight sealing of the ear canal, a foam earplug is inserted into the ear (Etymotic Research Disposable eartip ER1-14A, 13mm diameter). Before insertion users firmly squeeze the tip, which then expands within the ear to create a tight seal. To increase air-tightness further and to keep the electronic components in place, the 3D-printed case is sealed by applying hot glue on the backside of the components. The manufacturing process is the same for left and right earbuds, except that the device uses stereo sound which plays the respective channel on either of the ears. The pressure sensors in the earphones connect to an ESP32 MCU breakout board using I2C and data is transferred to a PC using serial communication. The audio signal connects to the same workstation using the aux connector.

5.2.2.3 Recognition Pipeline. Figure 5.4 illustrates the final recognition pipeline of EarRumble. Other classifiers and their performance are evaluated in subsection 5.2.4.

Detection. Activity detection uses a 360 ms sliding window with step size of 120 ms to decide if a window of the pressure signal contains activity. The detector computes the sum of absolute difference (SAD) within each window, and those with an SAD above 20 Pascal (PA) are flagged as containing activity. The threshold value was deduced from different rumble activity and non-activity samples that were collected during subsection 5.2.4, and ensures that $> 95\%$ of all samples are correctly detected. A correct detection is defined as one in which more than 75% of the gesture was detected as activity. An activity is considered complete and passed to the classifier after four consecutive windows without activity.

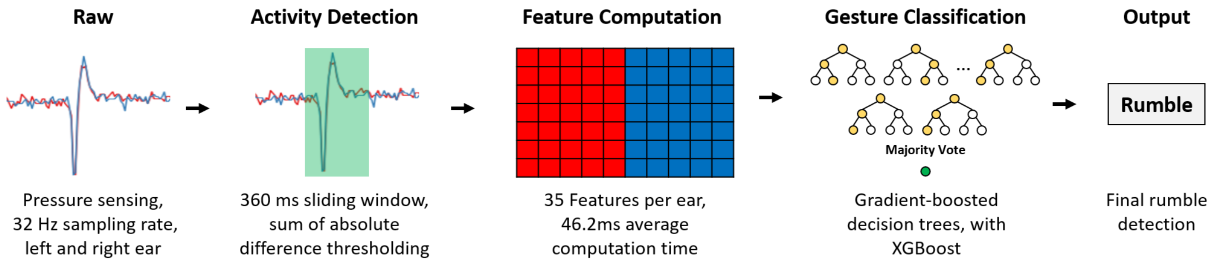


Figure 5.4.: The EarRumble recognition pipeline – the raw pressure signals are sampled at 32 Hz from both ears. A sliding window of 360 ms is used to detect if activity is present in the signal. 35 features are sampled from each ear, which are then passed to a gesture classifier to identify if the signal contains an ear rumble.

Classification. Once an activity is detected, it is passed to the classification stage to determine whether an ear rumble has been detected. Features are extracted from the signal, before being passed to a gesture classifier. Four different classifiers were evaluated: a radial basis function (RBF) kernel support vector machine (SVM), k-nearest neighbours (kNN), random forest (RF), and gradient boosting (XGBoost). The XGBoost classifier yields the highest overall accuracy at 0.95. More details are presented in subsection 5.2.4.

For classification features were self-defined, and also derived from the work by Ando et al. [17], in addition to systematically selecting features from tsfresh [103] using the data gathered in subsection 5.2.4. All features are computed on the zero mean-shifted signal to account for drift in the pressure readings. To reduce the initial collection of 1,618 features systematically, importance selection with XGBoost feature importance scores was applied. If a feature was deemed relevant for the left or right ear only, it is computed for both ears to account for different laterality conditions. This results in a set of 70 features (35 per ear) for feature extraction, which are then passed to the classifier. The average computation time of all features across all collected samples was 46.2 ms (Intel Core i7-9700KF 8 x 3.7 GHz).

For replicability, the features are: the number of, and the mean distance between peaks (both negative and positive), the absolute difference of the first two peaks, and the minimum and maximum values of the signal and also their locations. Additionally, the absolute difference and ratio of maximum to mean, and also the minimum and maximum slope and intercept over a five sample rolling window. Also, the sum of absolute differences (SAD) of the whole signal and also SAD of four even sequences that each sample is cut into. Finally, autocorrelation of the signal (lag: 2, 5, 9, 32), the variance of quantile changes (ql - qh: 0.2 - 1.0, 0.0 - 0.8), the spectral welch density (c: 2, 5), the continuous wavelet transform of the Ricker wavelet (c / w: 0 / 20, 2 / 2, 4 / 5, 8 / 20) and the FFT coefficients of the signal (c: 1, 2).

5.2.3 Prevalence of Tensor Tympani Muscle Control

Although not part of the acoustic reflex, the involuntary contraction of the tensor tympani muscle (tympani reflex) helps prevent ear damage from loud noises [218]. However, not everyone can voluntarily contract the muscle to cause in-ear vibrations, and there is currently no data about the prevalence of ear rumbling, nor how well people can voluntarily control the tensor tympani muscle. In this section, a large sample of participants was surveyed remotely using an online questionnaire, predicated on the basis that voluntary contraction of the tensor tympani can be self-reported due to audio feedback during the contraction of the muscle.

The dissertation aims to gain insight into what proportion of the population can voluntarily control the tensor tympani muscle - to inform how many people could hope to use ear rumbling for interaction. Of those who can voluntarily control contraction of the tensor tympani, deeper insight should be gained with respect to the level of control, isolation, and laterality when performing the ear rumbling. In the questionnaire, the level of control that participants have when contracting the muscle was queried to inform the potential complexity of the interaction design space. The questionnaire also investigates the level of discreetness afforded by ear rumbling as an interaction technique, which is dependent upon whether the tensor tympani can be contracted in isolation of other physical movements or facial gestures. Finally, this research seeks to discover whether participants can perform ear rumbling in one ear or both, which informs whether signals from the ears should be treated independently or in combination.

5.2.3.1 Design and Procedure. To minimise non-response bias – where those who can not ear rumble find the questionnaire less appealing – participants were recruited using neutral, context-free online ads and a social media post (Twitter) that did not reveal information on the nature of the study. Participants could use any device with browsing capabilities to fill in the survey, and no reward was offered for participating in the study. Firstly, participants were presented with the information sheet and relevant consent forms. Participants were not admitted to the study if they self-reported acute ear-related health conditions or were not at least 18 years of age.

The first page introduced the concept of "ear rumbling" by describing the contraction of the tensor tympani muscle. This was illustrated using an animated image of the human ear and a short textual description based on observations from people who have voluntary control, related work reporting the phenomenon, and by talking to an ear, nose, and throat doctor. Participants were then instructed to move to a quiet environment and to remove headphones, with explicit instructions for those who wear hearing aids to leave them on. Participants were asked to attempt to activate, flex, or move the muscles inside their ear to produce a rumbling sound. In addition, it was clarified that this is not to be confused with ear wiggling, and that some people can only perform ear rumbling when performing other actions (e.g. when yawning, swallowing, with closed eyes, or with the mouth open). Participants were then asked whether they could hear a "rumble or vibration", or a "clicking, crackling or popping" sound, or both. The question about the clicking, crackling, or popping sound was asked because such sounds might be induced by opening the Eustachian tube for pressure exchange [306], rather than contraction of the tensor tympani.

Participants who reported that they can perform ear rumbling, either with or without crackling noises in addition, were asked which ear they could hear the rumbling in and whether they could perform the rumbling independently of other actions, such as closing the eyes, blinking, or swallowing. They were also asked to complete two 7-point Likert items (1: strongly disagree, 7: strongly agree), one asking whether the rumbling is easy to perform on demand, the other asking whether it is comfortable to perform. To investigate the level of control users have of contracting the muscle, participants were asked whether they could perform the rumbling in quick succession (e.g. one ear rumble directly after another), and whether or not they could control the duration of the rumbling (e.g. hold an ear rumble for one second). The order in which these questions were presented was counterbalanced in the event that performing one movement made it more difficult to perform the other. If participants answered yes to either of these, they were asked the same two Likert items regarding ease of performance on demand and comfort, in addition to whether any additional action was required for the rumbling. The whole survey took around seven minutes to complete.

To validate the results from the online questionnaire, the same questionnaire was given to sixteen participants who completed the questionnaire as part of an in-person evaluation in subsection 5.2.4. These responses were validated visually using a USB otoscope and form a separate dataset. The responses to the Likert items from the online participants with those from subsection 5.2.4 were compared to see if any statistically significant differences exist.

5.2.3.2 Results. 208 participants completed the study, from a total of 1,399 clicks on the adverts. After data cleaning, there were 192 completed data sets (110 male, 78 female, 1 other, 4 preferred not to answer, age: M = 40.1, SD = 13.7, min = 18, max = 76). Eight participants had reduced hearing abilities (4 medical and 4 self-diagnosed), three wore hearing aids on both ears, and one participant was deaf.

Prevalence. Out of 192 participants, 83 (43.2%) reported that they could produce a rumbling or vibrating sound on at least one ear. Using the normal approximation interval, this results in a 95% confidence interval between 36.2% to 50.2%. Of the 83 who reported that they could produce a rumbling, 18 reported that they heard a crackling, clicking or popping sound in addition to the rumbling. Out of the 18 participants who reported rumbling and popping sounds, 9 said that they can do the rumbling sound independent of the crackling. In addition to those who reported some form of rumbling, 44 participants reported a crackling sound, but no rumbling. Those who did not report a rumbling sound were not asked to complete the remainder of the study. The deaf participant did not report rumbling, however participants with reduced or no hearing ability might still be able to contract the tensor tympani, but are lacking the audible feedback loop.

Laterality. Of those who could rumble, 68 participants (81.9%) reported the rumbling sound on both ears simultaneously, 11 (13.3%) in isolation on the left, and 19 (22.9%) in isolation on the right ear. Overall, 14 (7.1%) participants reported that they could perform rumbling on both ears and also in isolation on one ear - suggesting a high level of control of the muscle contraction. No participants reported the ability to perform rumbling on both ears and in isolation on both sides.

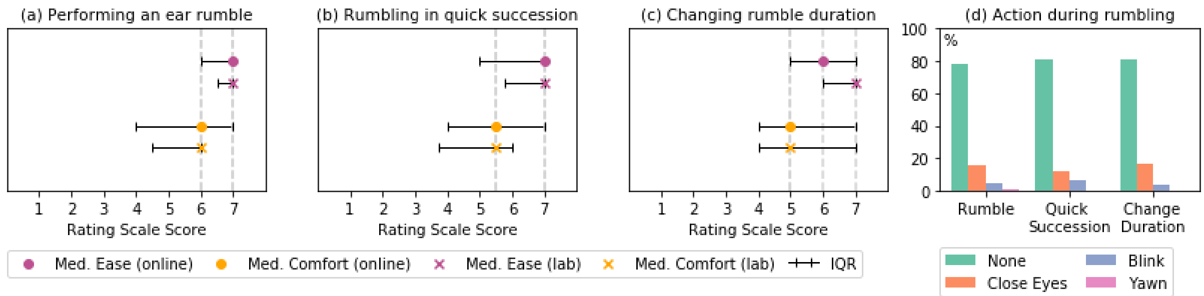


Figure 5.5.: The median and inter-quartile range (IQR) for online dataset from subsection 5.2.3 and lab participants from subsection 5.2.4 for the two Likert items – ease to perform on demand and comfort for performing – for (a) a rumble, (b) rumbling in quick succession, and (c) changing the duration of rumbles (1: strongly disagree SD, 7: strongly agree SA). Figure (d) shows the fractions of required secondary movement during different rumbling variations.

Control. Out of all those who could rumble, 74 (89.2%) reported the ability to perform rumbles in quick succession, and 71 participants (85.5%) could change the duration of rumbles. 67 participants reported that they could both perform rumbles in quick succession and change the duration of the rumbles. Interestingly, three of the participants with reduced hearing abilities could perform all rumbling variations (1 medical, 2 self-diagnosed). Figure 5.5 (a)-(c) shows that participants perceived ear rumbling to be easy to perform on-demand and comfortable. Performing a Friedman test for those who reported that they could perform all types of ear rumble, revealed a significant difference for responses about how easy ear rumbling is to perform on-demand, but no difference for the reported level of comfort across the rumble variations. Posthoc Wilcoxon rank-sign tests revealed that performing a single ear rumble was perceived to be much easier to perform on demand compared with both successive rumbles ($Z=37.5$, $p < .01$), and when holding a rumble for a prolonged duration ($Z=27.5$, $p < .01$). There was no statistically significant difference between successive rumbling or holding for a prolonged duration.

Isolation. The majority of participants did not have to perform secondary actions when contracting the tensor tympani muscle. Five participants required a secondary action when performing repetitive rumbling, and six participants when holding the rumbles for a longer duration. Figure 5.5 (d) shows the breakdown of the three types of secondary action which were required by these participants: closing eyes, blinking, and yawning.

Comparison to Lab Dataset. To validate the reliability of the Likert items administered online, the responses from the online participants were compared with the responses from participants who completed the same questionnaire in-person as part of the data collection performed in subsection 5.2.4. Mann-Whitney U tests showed that no significant differences existed between the responses across the three types of rumbling (single, changing duration, and successive rumbling) for either of the Likert items. Both data sets showed that participants found ear rumbling easy to perform on-demand, and comfortable, see Figure 5.5.

Qualitative Results. Five participants commented that they remember being able to perform ear rumbling since they were a child (P492, P556, P1320, P1809, P1709) and were surprised to learn that not everyone has that ability. Another participant said that they can recall involuntary rumbling during stress (P1320). One participant expressed the urge to keep rumbling after rumbling once (P480). Another participant with mild tinnitus mentioned that the rumbling can only be perceived when the surrounding area is very quiet. One participant expressed that rumbling led to a tense feeling in the cheek muscles.

5.2.3.3 Discussion. These results demonstrate that a substantial proportion of those who responded had the ability to voluntarily contract the tensor tympani, and generally participants reported that ear rumbling is easy to perform on demand and comfortable. It is unclear if and how the muscle contraction might be learnable for those who do not possess the ability to voluntarily contract the tensor tympani, or indeed if the muscle can be strengthened over time for those who can. Very few people required secondary actions (such as closing the eyes) to induce the ear rumble, which makes it ideal as a discreet interaction technique that is externally hard to notice. Of those who reported they could perform an ear rumble, 80.7% reported that they could change the duration of the ear rumble, and perform multiple ear rumbles in succession. Using this knowledge, it can be investigated how a simple gesture set that leverages these characteristics can provide simple input capabilities using contraction of the tensor tympani. The next section explores this in more depth.

The results reported in this section are predicated on the ability of participants to accurately self-report the ability to voluntarily contract the tensor tympani. In subsection 5.2.4, all participants (N=18) were initially recruited based on self-reporting the ability to ear rumble, which was validated using a camera otoscope and confirmed in all cases. It is also noted that the subjective responses of the in-person participants were not significantly different compared with the online survey. However, due to the nature of the online survey it was not possible to physiologically validate the ability to voluntarily contract the tensor tympani for the remote participants, and it is important to note that the data collected does not stand as physiological evidence.

5.2.4 Data Collection and Analysis

Next, a study was conducted to investigate the feasibility of detecting ear rumbling across a range of participants using in-ear barometry. Through this, the dissertation seeks to understand how well users can perform ear rumbling on-demand, how comfortable they find it, and the general characteristics when performing ear rumbles to inform interaction design. Based on initial data exploration, the data collection focuses on the use of three ear rumble gestures: single, double, and hold rumble.

The study also investigated participants' ability to perform repetitive sequential rumbling, where they were asked to contract the tensor tympani multiple times in quick succession. To goal of this task was to understand the level of control participants had over the muscle contraction, because a higher level of control opens up the opportunities for the use of switch scanning interfaces [425], rhythmic patterns [155], or beat synchronization [467; 504] for interaction.

Previous work has demonstrated that tensor tympani contraction might occur during swallowing and vocalization [245], and that in-ear pressure sensing can be used to detect a number of different gestures that users may naturally perform, such as opening and closing their mouth [17]. This begs the question as to whether the recognition pipeline can accurately distinguish pressure changes from ear rumbling compared with similar everyday activities. Therefore, the following actions which may induce pressure changes in the ear canal are also collected: opening and closing the mouth, reading out loud, drinking and swallowing, and chewing gum. The tensor tympani muscle contraction may also be induced by sound [122], therefore ear rumbling is investigated under two conditions: in silence, and with music playing through the earphones.

5.2.4.1 Participants and Apparatus. 18 participants were recruited through e-mail and a university Facebook group. Prior to the study, participants self-reported that they could perform ear rumbling, which was validated using an otoscope and confirmed in all cases. The study was conducted according to national COVID-19 regulations and within the university's safety guidelines. Participants wore the custom-built, in-ear pressure sensing device on both ears, see subsubsection 5.2.2.2. Each participant's outer ear canal was measured using a caliper, and two participants were excluded because of insufficient sealing of the ear canal with the ear buds (> 14 mm external ear canal diameter). The final dataset consists of 16 participants (13 male, 3 female, Age: M = 24.7 SD = 2.63, Ear canal width: M = 8.0 mm SD = 1.6 mm, Ear canal height: M = 11.2 mm SD = 1.8 mm). All participants reported that they had no hearing loss, and none of the participants wore a hearing aid.

5.2.4.2 Design and Procedure. Participants began the study by completing the same questionnaire about their ability to ear rumble featured in subsection 5.2.3, and all participants reported that they had not participated in the online questionnaire in subsection 5.2.3. Following this, the researcher measured the ear canal and verified that the ear rumbling reported by participants was caused by contraction of the tensor tympani muscle. This was validated by visual inspection of the eardrum with a Teslong USB digital in-ear camera otoscope. Participants were then asked to wear the EarRumble earphones and ensure a tight fit so that pressure differences could be detected.

Participants performed nine activities – four ear rumbling gestures (single, double, hold, repetitive), four everyday activities (opening/closing mouth, reading out loud, drinking, and chewing gum), in addition to an activity where users were asked to do nothing. A display was used to indicate which activity the participant should perform. Each activity was preceded by a five second on-screen countdown, after which participants were instructed to execute the activity in a five second window. Participants were asked to perform the rumbling gestures immediately and as quickly as possible after the countdown. Data was recorded from the start of the countdown, until the five seconds of the activity had elapsed. After each activity there was a five second break. After performing one of the rumbling activities, participants completed two 7-point Likert items (1: strongly disagree to 7: strongly agree):

- *Ease-on-demand:* The [rumbling activity] was easy to perform on demand.

- *Comfort* The [rumbling activity] was comfortable to perform.

The study follows a within-subject design, where all participants performed all activities. Participants performed the activities in two blocks: one in silence, and one with music playing (Symphony No. 5 by Ludwig van Beethoven). The order of the blocks was counterbalanced, i.e. half the participants performed in silence and then with music, and half vice versa. For each block, participants were asked to perform all rumble gestures, prior to performing all of the everyday activity tasks in a fixed order: open/close mouth, vocalization, drink and swallow, chew gum, and finally do nothing. The order in which the rumble activities are presented was counterbalanced across participants using a balanced Latin square design. For each activity, participants had a training phase prior to the data recording, in which each activity was repeated five times. The study lasted approximately 60 minutes, and participants received a bag of candy as a reward for their participation.

5.2.4.3 Data Labelling. The collected data was labeled by hand to be used for analysis and development of the EarRumble recognition pipeline. One researcher precisely identified the start and end time of single and double rumbles, and a second researcher verified the labels. The start time was labelled for the hold rumbles, however the exact end time could not be identified from the recorded data for the majority of samples because it was observed that the pressure equalised over time and there were no visible features to identify the end. In total there were 79 hold rumbles that were labelled by hand (49.3%). For repetitive and double rumbling, the peaks of each rumble were labeled to extract the periodicity. The peak detection was automated, using the Python Scipy signal processing library, and hyperparameters were fine tuned if they did not fit the individual samples. The accuracy of the peak detection was verified by visual inspection for all samples.

5.2.4.4 Results. The following section explores the participants' perception of ear rumbling and the characteristics of how users contract the tensor tympani muscle when performing different gestures. Figure 5.6 shows samples for all the different activities that participants performed during the study. All statistical results are reported as significant if $p < .05$, unless using Bonferroni correction to account for multiple comparisons.

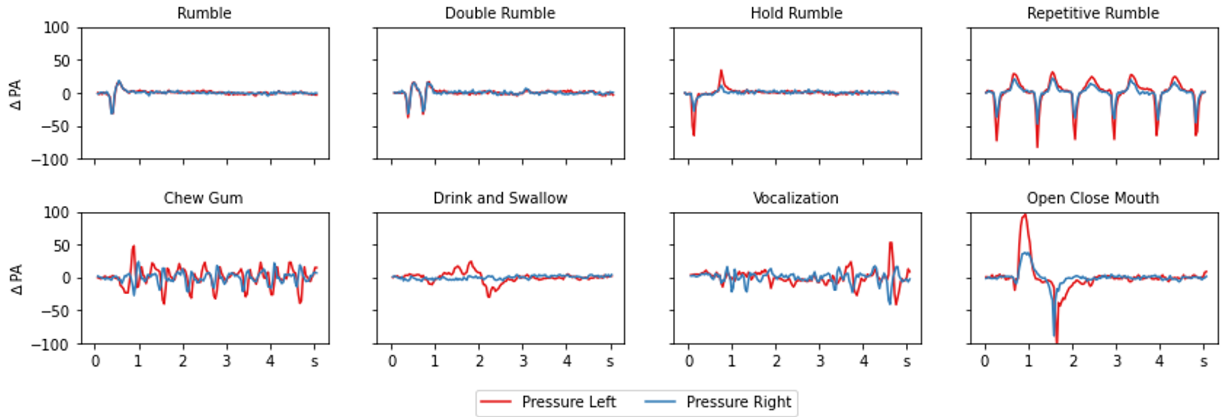


Figure 5.6.: Zero-shifted pressure readings measured in both ears for the different activities. The four rumbling variations show how the pressure readings spike downwards with initial muscle contraction and indicate another peak in the opposite direction after relaxation – creating distinct patterns for rumbling variations. The four noise activities have different characteristics.

Questionnaire Responses. First, the responses to the questionnaire that was also administered in subsection 5.2.3 are investigated. All participants correctly reported that they could rumble, which was validated visually with an otoscope. Five participants self-reported that they heard crackling in addition to the rumbling. Fourteen people reported that they could rumble in both ears, with the remaining two reporting that they could only rumble in their right ear. Three people out of the fourteen reported that they could perform ear rumbling in isolation in one ear in addition to both ears (1 right, 2 left). No participants reported that they could perform ear rumbling in both ears and in isolation in both left and right ears. Visual inspection of the pressure sensor data confirmed the reported laterality, with little to no pressure changes being observed when participants reported rumbling in only one ear. All participants reported that they could change the duration of the ear rumble, and that they could repeat the rumble in quick succession. Similarly, all participants reported that they could contract the tensor tympani without other activities, such as closing their eyes.

User Perception. Figure 5.7 shows the median and inter-quartile ranges for the Likert items participants completed during the data collection (please see Figure 5.5 for the Likert responses for the lab participants' responses to the questionnaire from subsection 5.2.3). Wilcoxon signed-rank tests showed that there was no significant differences for responses between the music or silence conditions across all the different rumble gestures. In general, participants reported that the ear rumbling gestures were easy to perform on demand, and comfortable. A Friedman test on the "easy to perform on demand" and "comfortable" Likert items was performed to see whether participants' perception was consistent across the different types of ear rumbling gestures. Participant results were significantly different for both the easy to perform ($\chi^2(3) = 11.79, p = .008$), and comfortable Likert items ($\chi^2(3) = 10.22, p = .017$). Pairwise comparisons using the Wilcoxon signed-rank test were performed with a Bonferroni correction for multiple comparisons. These revealed that participants found it significantly easier to perform a single rumble compared with holding the rumble for approximately 1 second (Z = 4.0, p = 0.0079). No other results were significant after Bonferonni correction.

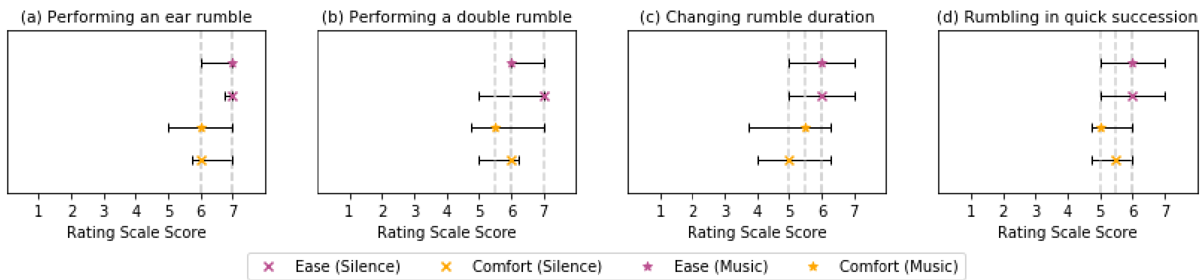


Figure 5.7.: Likert item responses from the 16 lab participants after performing rumble gestures in silence and with music.

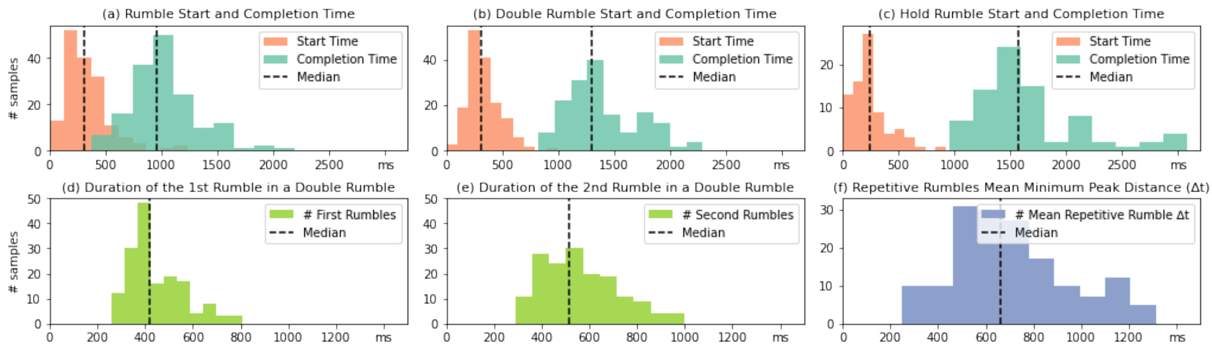


Figure 5.8.: (a)-(c) show the distribution of start and completion time for the different rumbling variations. Interestingly, a double rumble does not take twice as long as a single rumble whereas figure (d) and (e) show how the first rumble is faster to execute than the second. (f) shows how for repetitive rumbling the mean cycle time was significantly longer than during double rumbling.

Analysis of Ear Rumble Gestures. Figure 5.8 (a)-(c) shows the time to start the ear rumble when performing each of the three different ear rumble gestures (rumble, double rumble, hold rumble). The start time incorporates the participants' reaction time to the visual stimulus, and the time taken to start the gesture. As participants were presented a five second countdown, the reaction time should be smaller than that of a random stimulus. Statistical tests were performed to see if there was a significant difference between the reaction times across gestures, however due to the difficulty in extracting ground truth labels for the hold rumble, only single versus double rumbles were compared. Shapiro-wilks test of normality revealed both distributions were not normally distributed ($p < .05$), and hence the Wilcoxon signed-rank test was used. There was no significant difference between the start time for the single (Median = 308 ms) compared with the double rumble (Median = 308 ms). A Wilcoxon signed-rank test ($Z = 407.5$, $p < .001$) showed that a single rumble (Median = 958 ms) was significantly quicker to perform than a double rumble (Median = 1301 ms) - as to be expected. Interestingly, when analysing the duration of each rumble gesture, it was observed how the double rumble (Median: 1010 ms) takes only 48% longer than a single rumble (Median: 684 ms). Figure 5.8 (d) and (e) shows a histogram of the durations for the first and second rumble in the double rumble gesture. By comparing the time difference between rumbles, it can be seen how the first rumble when performed in the double rumble gesture is performed significantly quicker (Median = 419 ms) than the second (Median = 516 ms), using a Wilcoxon signed-rank test ($Z = 425.0$, $p < .001$) as both distributions were not normally distributed according to the Shaprio-Wilks test. Furthermore, it was found that the first ($Z = 313.0$, $p < .001$) and second rumbles ($Z = 2237.0$, $p < .001$) were performed significantly more quickly compared with the duration of the single rumble. Figure 5.8 (e) shows the average time between rumbles for the consecutive rumbling condition. In contrast to the double rumble, the time difference between rumbles when participants were asked to perform for 5 seconds was significantly longer (Median = 662 ms), and more comparable to the duration of a single rumble. A Wilcoxon signed-rank test showed there was no statistically significant difference between the single rumble and those performed in repetition.

Table 5.1.: Performance metrics comparison of different classifiers. XGBoost yielded the best overall performance for the individual rumbling activities.

Classifier	**Pre.**	**Rec.**	**F1**	**Acc.**
Dummy	0.28	0.28	0.28	0.28
RBF Kernel SVM	0.81	0.79	0.80	0.79
kNN	0.82	0.81	0.81	0.81
Random Forest	0.95	0.95	0.95	0.95
XGBoost	**0.95**	**0.95**	**0.95**	**0.95**

Classifier Performance Evaluation. To assess the performance of different classifiers under ideal conditions, the labeled start and end times of the different rumbles and time-constrained noise activities were used. The continuous activities were randomly sub-sampled from the recording (1 - 3 seconds). To avoid over-fitting the noise class, they were randomly selected evenly per class from all activities. For each classifier candidate, a 5-fold nested cross-validation with grid search for hyper-parameter optimization was performed. Both silent and music conditions were included in the training and test sets.

The best classifier was the XGBoost model which achieved 95% overall accuracy. The optimal hyper-parameters to achieve this were learning rate (0.08), maximum depth (3), and number of estimators (640). Table 5.1 shows the performance metrics of the different classifiers. The random forest classifier achieves similarly good performance overall, but it was found that it did not perform as well on the single rumble class.

Figure 5.9 shows the confusion matrix of the best classifier. Prolonged and double rumble achieve the best results. The main reason for the confusion between rumbles and noise classes is that rumbles are much less significant in their structure and therefore might be confused easier with short noise samples. Likely, with additional sensors (e.g. IMU or microphone) the noise classes could be discriminated with higher accuracy from the rumble class and vice-versa.

True Label \ Predicted Label	Rumble	Double Rumble	Prolonged Rumble	Noise
Rumble	0.86	0.04	0.03	0.07
Double Rumble	0.01	0.97	0.01	0.01
Prolonged Rumble	0.00	0.01	0.98	0.01
Noise	0.05	0.01	0.01	0.93

Figure 5.9.: The conufsion matrix shows how short, sub-sampled noise activities can confuse the classifier for single rumble detection and vice versa.

Leave-One Subject Out Validation. There may be variations in the data for the different rumble gestures between participants. The variability could either be temporal (e.g. completion time of rumbles) and/or due to differences in the intensity of the ear rumble (i.e. peak amplitude of the rumbles). Therefore, the best classifier was trained in a leave-one-subject-out cross-validation setting. Overall the classifier achieved 93% overall accuracy (2% decrease), which suggests that the proposed selection of features generalizes well across participants.

5.2.4.5 Discussion. It was demonstrated how in-ear barometry can be used to detect contraction of the tensor tympani. The hold rumble gesture was particularly difficult to identify ground truth labels for, and this inherently has ramifications for the development of a classification pipeline because there are fewer samples for training and validation for the hold rumble. It also implies that a large percentage of hold rumbles may not be accurately detected using in-ear barometry. However, participants found contraction of the tensor tympani easy to perform on-demand and comfortable across the different ear rumble gestures.

Ear rumbling can be detected in silence or with music playing, however one participant mentioned that it was harder to focus on the execution of rumbles because the music made the rumbling sound harder to hear which affected the feedback loop. This is something that was potentially also observed in the previous online questionnaire with the deaf participant and the participant with mild tinnitus, as feedback of the gesture is an important part of the interaction to notify that their interaction has been successfully registered [50]. However, it also important to note that in this context there was no action associated with the ear rumbling gestures, and feedback can be provided either indirectly through the response of the system (e.g. changing the song or answering the phone), or directly in response to detection of the ear rumbling itself (e.g. play a sound to indicate rumbling).

There are a number of interesting insights to be gained from the analysis of rumbling characteristics in Figure 5.2.4.4. The acquisition time of an input technique refers to the time required to acquire the input device so that it is ready for use (e.g. unsheathing a pen), and the homing time refers to the time required to return to a "home" position (e.g. making contact with a finger for touch screen interaction) [192]. Ear rumbling through contraction of the tensor tympani does not introduce any acquisition time, nor any homing time. During the data collection participants were presented with a countdown timer prior to performing the ear rumble gestures, therefore it can be seen how quickly users are able to respond to the predicted stimulus.

No significant difference were observed for the start times between performing a single or double rumble, however it should be noted that it takes approximately 308 ms to begin the ear rumble. This may be due to the fact that voluntary contraction of the tensor tympani is rarely performed as it serves little purpose. Comparing the time to start a rumble gesture, it is nearly 100 ms shorter than to home in on a device (400 ms) according to the keystroke-level model (KLM) [87]. Based on the duration of a single ear rumble it can be noted how it is similar to typing random letters (500 ms) or complex codes (750 ms) according to KLM.

The insights into repetitive rumbling indicate that contraction of the tensor tympani could be suitable for more complex interactions than the three basic gestures that were explored here, which opens up the opportunity to use ear rumbling with switch scanning interfaces [425], rhythmic patterns [155], or beat synchronisation [467; 504]. The results provide insights into what kind of tempo one could use for these interactions to optimise throughput of the technique, and participants reported that this was generally easy to perform on-demand and comfortable. There is also scope to incorporate ear rumbling interaction with existing techniques, such as extending Ando et al.'s in-ear barometry-based gesture set [17], so that a wider vocabulary is available for users when interactions beyond simple binary choices are required.

5.2.5 Usability Evaluation

A study was performed to explore the performance and usability of EarRumble as a hands- and eyes-free input technique. Using a real-time implementation of the pipeline featured in subsubsection 5.2.2.3, the exploration is grounded in three manual, dual task application scenarios that one might face when using earables in everyday life. The goals of the study was to test how well the pipeline worked, and to gather feedback from users when using EarRumble for interaction.

5.2.5.1 Participants and Apparatus. Eight participants (7 male, 1 female, Age: M = 24.5 SD = 2.7) were invited from the previous data collection study in subsection 5.2.4. Participants were seated in a regular office chair in front of a desktop PC. The pressure sensors in the earphones were connected via USB to a desktop computer which was running the EarRumble detection pipeline software in Python. Participants used the custom-built EarRumble earphones described in subsection 5.2.2, and a separate pair of HolyHigh in-ear earables which features a mechanical click button on the outside of both earphones. Participants were instructed to use either earphone for the click, depending on their handedness preference.

5.2.5.2 Design. Tree rumble gestures (single rumble, double rumble, and hold rumble) using the EarRumble technique were compared with a simple button click on a pair of smart earphones using three analogous gestures (click, double click, hold). A button click was chosen instead of a tap gesture because the button allows for the accurate detection of hold gestures. Three different use cases using two applications were selected for the evaluation. These were chosen to evaluate the techniques in the context in which they may be used in real-life.

Incoming Call. The first scenario featured an incoming call, whereby a ringtone plays in the earphones and the user can either accept (rumble/click), reject (double rumble/click), or mute (hold rumble/button) the incoming phone call. A use case was mimicked in which the user's attention is on a manual task. Participants were given a primary task of typing a piece of text on the desktop PC whilst playing music in the background through the earphones. The phone would then ring and the user would be tasked with either accepting, rejecting, or muting the phone call, before carrying on with the typing task.

Audio Player. The second and third application scenarios feature an audio player in which the user can play and pause (rumble/click), skip to the next (double rumble/click), or go back to the previous track (hold rumble/button). The second scenario consists of an audio transcription task whereby participants are tasked with transcribing sentences being read aloud through the earphones. After each sentence the participant is required to pause the audio, write the sentence down, and resume playback to hear the next sentence. The third scenario consisted of a music playlist in which participants were required to skip forwards or backwards in order to find specific songs.

5.2.5.3 Procedure. Participants began the study by completing demographic information and signing a consent form. They were then given one of the application tasks to practice with, using both input techniques. The order in which the applications were presented to participants were counterbalanced, as was the order of the input techniques.

For the incoming call scenario, participants completed the task in three blocks. For each block, the participant had to accept/reject/mute all incoming calls. Only one action was chosen per block to reduce the burden of memorising which action to take. In each block, the participant received four phone calls at intervals of 20 seconds. This was chosen to give the participant time to resume typing, and long enough to reduce the chance of precisely predicting when a call would occur. For the audio player scenario, participants were given four sentences to transcribe, resulting in eight rumbles/clicks in order to pause and resume playback. The song playlist consisted of five songs, and participants were tasked with finding the songs which involved skipping 3x forward, 2x back, 1x forward, and 2x back – resulting in four gestures each of the double rumble/click and hold rumble/click.

For the incoming call task the response time of participants from when the call was triggered to the corresponding action was measured. Also the time taken to return to the typing task was measured, defined as the first keystroke after the incoming call had been actioned (e.g. rejected). Participants were asked to report any errors during the interactions arising from (a) incorrect detection of gestures (e.g. rumbles/clicks which aren't detected or which should have been), (b) incorrect classification of gestures (e.g. detection of a single rumble/click instead of a hold rumble/click), and (c) user error (i.e. performing the wrong gesture). Participants completed the NASA TLX [179] and were asked what they liked and disliked about each input technique for both of the applications. After both techniques had been performed for an application, participants were asked which they preferred using for the specific application, and at the end of the study they were asked which one they preferred overall.

5.2.5.4 Results. When asked about their overall preferences, six participants preferred the EarRumble technique, and two preferred the button click. The EarRumble technique was favoured because it required less effort (P2, P4, P6, P8), and was more comfortable (P5). P3 specifically highlighted the single ear rumble gesture was their preferred technique. The button click was preferred because it was more robust (P1), and P7 noted that they would have preferred the earphones if it was tap input rather than button click because the EarRumble was lacking immediate feedback. Issues with the reliably of detecting the hold rumble were discovered – only 56% of prolonged rumbles were detected correctly compared with 91% of single rumbles and 94% of double rumbles throughout all tasks and across participants. Friedman tests on the responses to the NASA TLX revealed no significant differences between the input technique and task conditions. In the following the feedback from the two application scenarios is discussed.

Phone Call Task. For the phone call task, seven participants highlighted the advantage of not having to take their hands off the keyboard, and ability to continue typing immediately. P6 perceived that it felt much quicker than the button to interact. Excluding erroneous detections, the mean time between the phone call and detection of gesture was 3.40 s for EarRumble, and 3.34 s for the button click. The mean time between detection and the first keystroke during the call task was 1.09 s for EarRumble and 1.51 s for the button click. No statistical tests were run due to the different numbers of successful detections and low participant numbers. Interestingly, P1 also noted how it was nice not to have to use voice – alluding to the social acceptability of the technique. However, six participants struggled using the hold rumble for interaction. P1, P5, and P7 reported that the hold rumble required additional concentration and was more uncomfortable to perform. P4 also commented on the latency of the rumble detection, and both P1 and P7 noted the lack of immediate feedback from the EarRumble techniques.

The participants also saw advantages of using the button for input, because it was clear how it worked and have used it before (P1, P7), provided immediate haptic feedback (P1, P4, P8), and worked reliably (P1, P3) with low latency (P1). However, 3 participants highlighted a disadvantage of having to take their hands off the keyboard to interact (P1, P2, P4) which they felt broke their workflow. Also, five participants felt the physical click button was hard to press and hurt the ear canal (P2, P3, P5, P6, P7). Two preferred the button click because of the immediate feedback (P1) and because it was more robust (P7).

Six out of the eight participants preferred the EarRumble technique for the phone call task, despite the technical issues with the hold rumble gesture. Participants preferred EarRumble because it required no extra movement (P2, P6), less time (P5), and allowed them to continue typing (P3, P8). P4 preferred the technique because "it felt magical" and "almost felt telepathic".

Audio Player Tasks. Participants gave similar feedback for the audio player tasks. The EarRumble technique was perceived to be faster because it does not require the use of the hands (P1, P3, P4, P8), and was low effort (P5). P6 described the technique as "much more practical" than the button click. Interestingly, two participants described the interactions with the audio player using EarRumble as "fun" (P7, P8). The disadvantages cited once again referred to the hold rumble (P1, P2, P5, P6, P7, P8), and higher latency of rumbling detection (P3, P4, P5).

Feedback was similar as well for the button click with the audio player, with it being described as a known technique (P5), very robust (P1), and immediate haptic feedback was an advantage (P1, P4, P7, P8). Participants described it as "annoying" to take the hands off the keyboard when pausing the text to write (P1, P3, P4), and once again they highlighted that the click hurts the ear canal (P2, P3, P5, P6, P7, P8).

No participants changed their preference between phone call task and audio player task. Participants preferred the EarRumble technique because it was low effort (P3, P4, P8), and because it does not require the use of the hands (P6). P2 said they preferred EarRumble because they think it would be "perfect when listening to music in bed". P5 reported that they preferred EarRumble because the button click was uncomfortable to use, and the remaining participants preferred the button because it was more robust (P1, P7).

5.2.5.5 Discussion. This usability evaluation highlights the low-effort, hands-free nature of EarRumble, which was the main motivation of adopting ear rumbling for interaction. Scenarios in which participants could imagine using the EarRumble technique included during focused work (P1, P2, P4), when hands are occupied (P2, P5, P8), for secretive input (P2, P3), to interact without any noise, e.g. speech (P3), or for music or calls (P6). However, as expected from the results in subsection 5.2.4, the pressure sensing technique did not reliably detect the hold rumble gestures. For some users, the accuracy was very high (90+%), however there are larger issues around detecting the hold rumble using in-ear barometry. It is also noted that the latency of the pipeline was an issue for some participants, which with further optimisation could be further reduced allowing for quicker selection times.

5.2.6 Future Work and Limitations

EarRumble requires air-tight sealing of the user's ear canal. Blocking the ear canal with headphones for prolonged time can change the "climate" of the ear (e.g. temperature and humidity) and is often said to support the entry of, e.g., bacteria in the middle ear. However, no significant clinical evidence exists to back increased bacterial or fungal exposure by continuous use of regular earphones [305; 106]. Regarding comfort, one participant mentioned that they felt uncomfortable sealing the ear canal for prolonged periods, and this was also applicable when wearing regular in-ear headphones (subsection 5.2.5-P7). The ear caps of the current EarRumble system are foam soft type plugs which fit in tightly in the ear canal and go deeper into the ear canal than regular ear plugs, meaning they may not feel as comfortable after expansion as regular in-ear type plastic caps. Initially, standard earphone plastic caps were used and still pressure changes could be observed clearly and consistently, however in the initial exploration of in-ear barometry it was noticed that some users could not perform the ear rumbles as strongly and they were harder to detect. Nevertheless, standard in-ear caps should be investigated further in future work for real-world applications.

Ear rumbling is a prime candidate for providing simple gestures on the go – for example, interaction in a crowded train would be easily possible without requiring any movement by the user, or other cases where mobility is limited. The use of ear rumbling was only investigated with users sat down in a stationary position, and did not investigate the social acceptability implications of the technique. Real-world deployments of the technology may reveal interesting insights into how movement in unconstrained environments affects the sealing of the ear canal, the detection pipeline, and/or a user's ability to contract the tensor tympani. It also may reveal how many other actions throughout the day could lead to false positives (e.g. yawning), as the false positive actions chosen in this dissertation were based on easily replicable actions that most closely resembled the contraction of the tensor tympani. There may also be scope to suppress false positives through the use of other sensing modalities, e.g. sensing chewing gum using an IMU [282].

The applied sensing principle allows for reliable detection of single and double rumbles, however the hold rumbles proved to be problematic and the current setup does not allow to derive the duration of rumbles precisely. In the future, other sensing principles may be used to realize ear rumble detection more reliably, e.g. by an in-ear camera or acoustic impedance measurements

[327; 326]. The latter might be even realized with off-the-shelf hardware with noise canceling earphones as they have an in-ear microphone to for the noise canceling feedback loop [12].

Finally, no literature was identified discussing the consequences of voluntary tensor tympani contraction over a long-term basis. Potential long-term safety concerns were discussed with the ear, nose, and throat doctor consulted during the project, who noted no known safety issues with voluntarily contracting the tensor tympani and could not see why this would cause any problems. However, the absence of data relating to this does not imply long-term safety, which future work should further investigate.

5.2.7 Conclusion

EarRumble uses in-ear barometry to detect the contraction of the tensor tympani muscle, known as ear rumbling, allowing users to provide low-effort, discreet interaction using earable devices. An online questionnaire showed that 44% of respondents reported that they could perform ear rumbling, and a data collection with 16 participants provided insights into the level of control users have over contracting the tensor tympani, demonstrating how ear rumbling is a viable interaction technique. The dissertation explored how interaction could be achieved using three simple "gestures" using a detection pipeline consisting of feature extraction and gradient boosted classification. Single and double rumbles could be accurately detected, however detection of a rumble that is held for a prolonged period (e.g. 1 second) proved to be problematic for many participants due to the pressure sensing approach. A usability evaluation grounded in three manual, dual task application scenarios showed the low-effort, hands-free advantages of the technique relative to providing input via a button on the earable. The use of ear rumbling has the potential to be useful in a number of application scenarios involving on-the-go mobile interaction, with scope for future work to investigate more robust sensing techniques.

6. Earables as Platform

This chapter presents contributions that go beyond the scope of specific applications areas and addresses the more general earable field and its fundamental technologies. Section 6.1 introduces an open-source hardware platform for earable computing called “OpenEarable” that integrates a series of advanced sensing capabilities into a modular, ear-worn device. This is followed by section 6.2 which introduces an evaluation of standard earable form factors with respect to their suitability to be worn during sleep, and an evaluation of eSense, another earable prototyping device, for problem-based learning.okay

6.1 Open Hardware for Earable Sensor Applications

Chapter 2 has demonstrated how earables integrate diverse sensing capabilities to detect a multitude of interesting phenomena that have been used in applications spanning the four research areas identified by this dissertation: physiological monitoring and health, movement and activity, interaction, and authentication and identification. As a result, the earable platform has attracted attention from several, different research communities, and the number of research publications using the platform is increasing year-on-year.

A wide variety of hardware prototypes have been used in the earable research literature, ranging from commercial offerings such as Apple Air Pods[1] and cosinuss[2], prototype research platforms such as *eSense* developed by Nokia Bell Labs [235], and fully bespoke earable research devices (e.g., [77]). Of particular note is the eSense, which has accelerated the growth of earable research within the academic community. This in part was driven by the devices being freely distributed to academics across the world, providing a platform for

[1] AirPods Pro - `https://www.apple.com/airpods-pro/`
[2] cosinuss - `https://www.cosinuss.com/en/technology/`

which earable research can place and which others can openly contribute to. More recently, Chatterjee et al. [94] introduced *ClearBuds*, which has open hardware and comes equipped with dual microphones that can be used for speaker separation using beamforming. However, these earable platforms lack the extensibility that is required to take full advantage of the wide range of sensors that have been shown to be effective on the earable platform.

Therefore, the dissertation introduces OpenEarable, the first open-source, Arduino-based earable research platform. OpenEarable aims to build upon the success of other earable prototyping platforms by providing a fully transparent and open hardware platform that enables researchers to push the boundaries of earable research. The main objective of OpenEarable is to provide an extensible platform that can be easily and cost-effectively manufactured for research and development purposes. OpenEarable features a 9-axis IMU (accelerometer, gyroscope and magnetometer), an ear canal pressure and temperature sensor, an inward facing ultrasonic microphone as well as a speaker, a push button, an on-off-switch, and a controllable LED. The following sections provide an overview of the design process and an in-depth walkthrough of the hardware and software systems that make up the OpenEarable platform. Based on three exemplar applications from the research literature it is highlighted how the platform has to the potential to be used for motion-based activity tracking, detection of chewing events, and ear canal shape based authentication. The presented work was published at EarComp 2022 [456].

6.1.1 Design Process

In the following, the guiding design principles and sensor selection for OpenEarable are rationalized.

6.1.1.1 Guiding Principles. The main objective when developing the OpenEarable platform was to provide a general-purpose hardware sensing platform for the earable research community that allows for the exploration of state-of-the-art sensing capabilities on the ear. This process was guided by the following principles throughout the design and development phase:

Openness and Extensibility. The OpenEarable platform's hardware and software should be open to, and easily extensible by, others. The OpenEarable platform should provide the core infrastructure to enable the exploration of different sensing paradigms. As a result, all hardware design files, firmware, communication interfaces, and data recording tools should be made public and easily accessible so that others can modify and expand the platform in unique and novel ways. In addition, the earpiece, which contains the critical sensor inside the ear canal, should be interchangeable from a core main processing component that deals with communication and processing. Also, a conscious effort was made to use development tools that are free-of-charge in the design of the hardware and software. This way, as many people as possible can be provided with the opportunity to develop on, and for, the OpenEarable platform.

Manufacturability and Cost-Effectiveness. In order for people to leverage the OpenEarable platform for research, they must be able to easily manufacture the device at an affordable cost. To achieve this, the system should use commercial off-the-shelf components that require no specialized tools for manufacture. The PCB was specifically designed to be manufactured, and components assembled, by a self-service PCB assembly manufacturer. Additionally, all plastic parts are designed to be 3D-printable with a standard stereolithography (SLA) printer, commonly available as consumer devices or available to order online. The assembly of an OpenEarable should only require minimal equipment, with the version presented in this dissertation only requiring a soldering iron, pliers and plastic-compatible power glue. All hardware components should be compatible with the open-source Arduino platform.

Attachment and Comfort. The OpenEarable platform, and any extensions, will need to be validated with users and therefore it should be easy to attach yet stable and robust against user movement. In addition, the earable should be comfortable to wear within the limitations of a general purpose prototyping device. Therefore, OpenEarable has an over-the-ear hook design that wraps around the auricle to encapsulate the electronics whilst providing mechanical stability. This provides an opportunity for sensors to be placed in, on, or around the ear via a flexible cable that runs from the main unit behind the ear to a sensing earpiece.

6.1.1.2 Sensor Selection. For the OpenEarable platform both traditional and new sensing capabilities not currently available on other earable platforms are incorporated. For basic input there is a push button, and a 9-axis inertial measurement unit (IMU) for motion-based applications (e.g., gait analysis [32]) and to filter out motion artifacts.

For new sensing capabilities an ultrasound microphone and an in-ear pressure sensor combined with temperate sensor were incorporated. Many earable platforms feature access to an external microphone for voice-based interaction, and most have a microphone inside the earbud for noise-cancellation. However, no available platform that provides access to a microphone placed inside the ear canal could be identified. Therefore, OpenEarable features an inward facing ultrasonic microphone which can be used to detect ear canal shape and deformation based on measured sound reflection. An ultrasound microphone was chosen to be able to detect both audible and inaudible sounds which do not disturb the wearer. Also, a pressure sensor was integrated for in-ear barometry which provides information about the ear canal shape and deformation. In-ear barometry has gained traction in recent years across a range of applications and have been used to detect jaw and facial movements [17], blood pressure [500], and contraction of the tensor tympani muscle (see section 5.2).

6.1.2 Hardware

The hardware of OpenEarable is inspired by existing works in the earable domain. The following sections present the electronics, mechanical design, and production process.

6.1.2.1 Electronics. The following section describes the circuit layout, microcontroller unit, power architecture and sensors of OpenEarable. A schematic system architecture overview is shown in Figure 6.1. The PCBs and other components for mechanical assembly are shown in Figure 6.2.

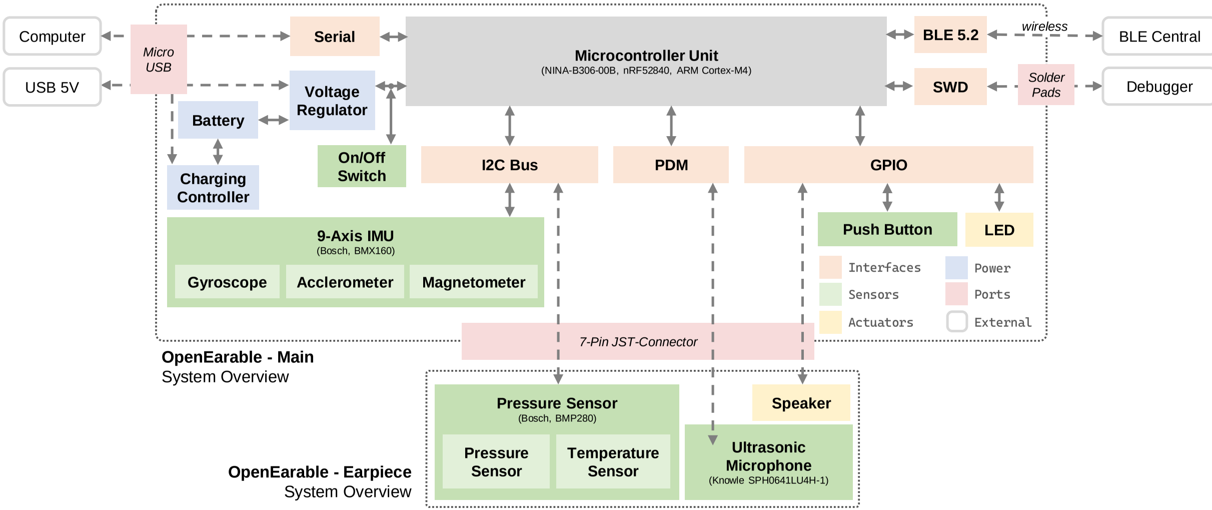


Figure 6.1.: System overview of the OpenEarable system architecture. The microcontroller unit is the central hub which communicates with sensors, actuators, and external devices. The earpiece is modularized into its own logical unit.

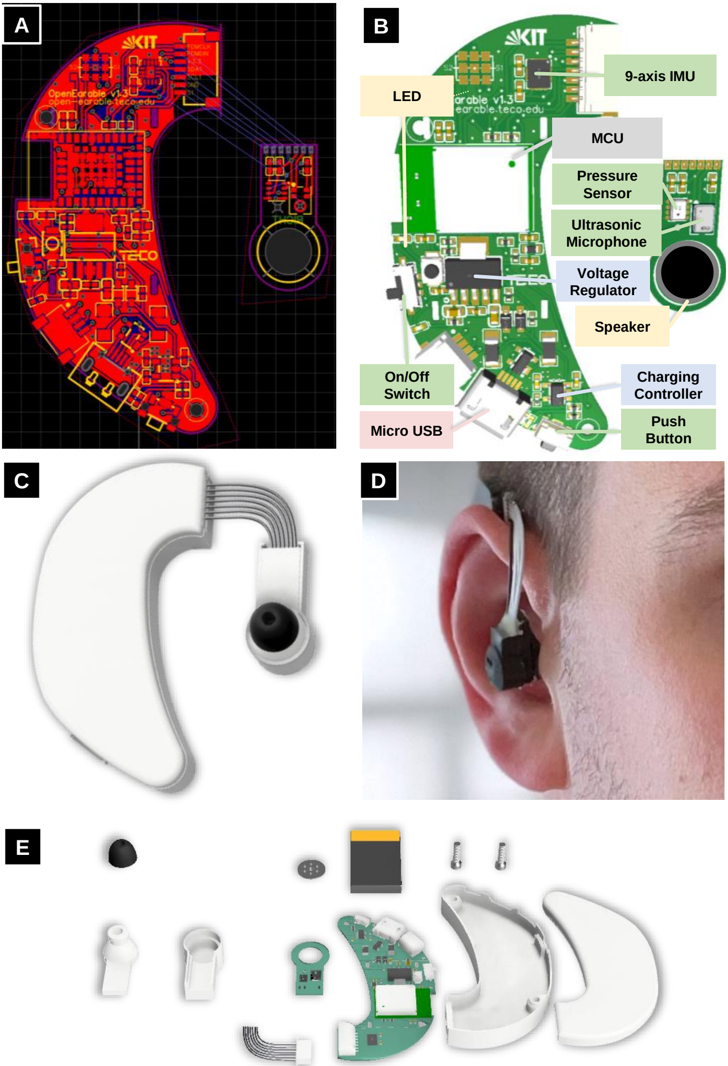


Figure 6.2.: (A) PCB layout of OpenEarable; (B) 3D-rendering of the PCB and components; (C) assembled device; (D) a person wearing the device; (E) disassembled hardware components.

Printed Circuit Board. OpenEarable consists of two PCBs: one that sits behind the ear to house all processing and the IMU as well as power management, and a sensing earpiece PCB (see Figure 6.1 and Figure 6.2 A, B , and E). The PCBs are connected via a 7-wire cable. The OpenEarable PCBs were designed in an ear-hook form factor which makes it easy to attach the device to the ear. In addition, the shape of the PCB creates sufficient space to place all components behind the ear comfortably. The PCBs are 0.4 mm thick and are designed so that all surface mount device (SMD) components are on the top side only which simplifies assembly and makes it possible to have the components placed and soldered by a self-service PCB assembly manufacturer. For the earpiece, there are two version: one for the left ear and one for the right with similar, mirrored components. The main PCB can be used on both sides interchangeably. Two holes in the earpiece PCB are designed specifically to let air and sound pass to the pressure sensor and ultrasonic microphone. A larger hole is in place for the speaker. Two screw holes were added to the main PCB to attach the 3D-printed enclosure. The design files of the PCBs are open-source and released under a CC-BY license.

Microcontroller Unit. The microcontroller unit (MCU) of OpenEarable is a u-blox NINA-B306-00B module which is based on the nRF52840 Bluetooth Low Energy (BLE) 5.2 system on chip (SoC). OpenEarable makes use of the Inter-Integrated Circuit (I2C) interface to communicate with the pressure and temperature sensor as well as accelerometer and gyroscope. The digital pulse density modulation (PDM) interface is used to read the microphone. Programming the MCU is possible using USB Serial or via Serial Wire Debug (SWD) on the backside of the PCB (e.g., to initially flash the USB Device Firmware Upgrade bootloader).

Power Architecture. In general, OpenEarable is intended to run from a single LiPo battery cell (*PowerCElls 301525, 90mAh nominal capacitiy, 3.7 V nominal voltage*). Charging is possible via a micro USB port with electrostatic discharge protection. For battery charging the board uses the *Microchip Technology MCP73831T* charging controller. As the MCU operates at 3.3V, OpenEarable also comes with a low dropout voltage regulator (*Texas Instru-*

ments TPS73733). It is possible to also use the device while charging. When sampling all sensors and sending out the data via Bluetooth Low Energy a fully charged OpenEarable lasts roughly 10 hours which is well above the threshold for most research .

Ultrasonic Microphone and Speaker. An ultrasonic microphone (*Knowles SPH0641LU4H-1*) with bottom port is placed in close proximity above the speaker. By default, OpenEarable samples the microphone at approximately 44 kHz. The speaker inside OpenEarable is a standard true wireless stereo (TWS) 8 mm, 16 Ohm resistance earbud component that is available from many consumer electronics stores.

Pressure and Temperature Sensor. Pressure and temperature are measured in close proximity to the speaker and ultrasonic microphone. A hole in the PCB next to the pressure sensors redirects airflow from inside the ear canal. The pressure and temperature information are available from a single package inside the *Bosch BMP280* pressure sensor. The sensor typically samples at 182 Hz in an absolute pressure range of 300 to 1100 hPa. The temperature sensor supports a similar sampling rate range and has an absolute accuracy of ±0.5 °C.

Accelerometer and Gyroscope. OpenEarable has a 9-axis inertial measurement unit (*Bosch BMX160*) compromising of a 3-axis digital accelerometer, a 3-axis digital gyroscope and a 3-axis digital mangetometer. Linear acceleration measurement range and angular measurement range can both be configured. In theory, the accelerometer supports 12.5 Hz to 1600 Hz, the gyroscope 25 Hz to 2500 Hz, and the magnetometer 12.5 Hz. Limited by BLE bandwidth, OpenEarable currently limits to 100 Hz.

Light Emitting Diodes. OpenEarable features four LEDs for basic output. Two static LED indicating the charging status when the micro USB cable is plugged in (red on, green off: charging, red off, green on: fully charged or not charging). One LED indicates when the power switch is turned on (green on: powered, green off: no power - i.e. switched off or battery dead). The fourth LED can be freely turned on and off or controlled in brightness using pulse-width modulation (PWM).

Push Button. A push button on the lower backside of OpenEarable can be used for simple, binary input. Another push button next to the MCU serves as reset button and can be used to enter the device firmware updates mode of the microcontroller by double pressing the button. An on-off-switch can be used to turn on and off the device.

6.1.2.2 Mechanical Design and Assembly. The assembly of the PCB was done by a contract manufacturer, see subsubsection 6.1.2.3. The remaining parts have to be self-assembled are described below.

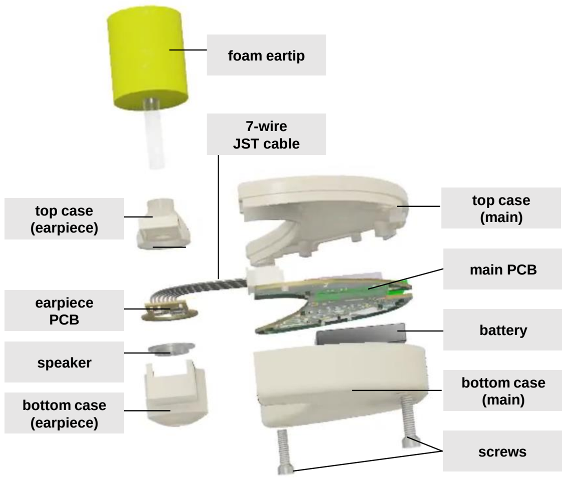


Figure 6.3.: (A) PCB layout of OpenEarable; (B) 3D-rendering of the PCB and components; (C) assembled device; (D) a person wearing the device; (E) disassembled hardware components.

Speaker and Battery. The speaker has an adhesive foam ring pre-installed so it can be glued onto the PCB while also sealing off the speaker (see Figure 6.3). In addition, it is held in place by the earpiece case. The battery is mechanically also held in place by the case and the wires are soldered ontp the PCB (see Figure 6.3).

Earplug. The OpenEarable earplug consists of two 3D-printed parts which are glued together and sealed off with plastic-friendly glue (Pattex instant glue). The front part sits above the speaker and the PCB through-hole and seals it of. Either, a foam type sealing earplug with plastic tube (*Etymotic Research* disposable eartip ER1-14A, 13mm diameter, see Figure 6.3 C (a)) to maximise ear canal sealing, or a silicone standard eartip can be put on the earplug (see Figure 6.3 and Figure 6.2). The backside separates the speaker cables and pressure sensor as well as microphone. Together, the 3D printed parts ensure that the ear canal is sealed for pressure sensing.

6.1.2.3 Production and Costs. OpenEarable was designed with the JLCPCB [3] parts library in mind. Therefore, almost all components are available as standard self-service SMD parts assembly order. The MCU and microphone have to be ordered specifically for assembly which, from past experience with JLCPCB, has two weeks lead time (depending on supplier availability) following which PCB manufacturing and assembly require an additional working week. For the 3D-printed parts an *Formlabs Form 3B+* printer was used, however, there are also many inexpensive online 3D-printing services available that could be used to manufacture the earpiece plastic parts made-to-order.

The total costs excluding shipping for ten OpenEarable is roughly \$550 ($\approx$ \$55 per device). The costs per device are split as follows: \$1 PCBs, \$43 electric components, \$4 foam earpieces (incl. 4 replacements), \$1 screws and \$4 3D-printed parts (if ordered online). One-sided PCB assembly is free of charge.

6.1.3 Software

All OpenEarable software is open-source and available on the project website under the MIT license.

[3] JLCPCB - `https://jlcpcb.com/`

6.1.3.1 Firmware. The OpenEarable firmware is implemented in C++ using the Arduino framework based on the implementation of the *Arduino Nano 33 Bluetooth Low Energy (BLE) Sense*. This makes it easily possible for others to change the firmware running on the device. The firmware reads out all sensors and makes them available via BLE. Due to bandwidth limitations, at least BLE 4.2 has to be supported by the device connecting to OpenEarable.

Generic Attribute Profile Specification. OpenEarable's main interface for data transfer is a custom-defined Generic ATTribute Profile (GATT). Based on the profile, various functionalities of the earable can be controlled as well as sensors can be configured and read out. Table 6.1 gives an overview of the GATT specification for OpenEarable for regular data recording, as well as for recording and sending audio data. The sensor service is responsible for enabling sensors, configuring sampling rates and sending out sensor data. Using the device info service, a unique name and the device generation can be read out. The dedicated audio service sends out bursts of audio samples of roughly 1 second duration sampled at 62.5 kHz. At the moment, continuous audio streaming is not supported, however this is a software limitation that will be fixed in a future iteration (see subsubsection 6.1.4.4).

Table 6.1.: BLE GATT profile services and characteristics overview of OpenEarable. A detailed documentation including UUIDs of the BLE API can be found on the project's website. The specification follows the schema for usage with edge-ml.org. (R = Read, W = Write, N = Notify)

Service	Characteristic	R/W/N	Description
sensor	sDatChar	R/N	timestamped sensor data
	sConfChar	W	enable and configure sensors
device	deviceIdChar	R	unique name of the device
	deviceGenChar	R	generation of the device
audio	audioChar	R/N	burst chunks of ultrasonic audio
	packageInfChar	R/N	package info and sending state

6.1.3.2 Recording Tool.

Two options are available to record data with OpenEarable, a custom-built dashboard and an open-source and browser-based toolchain for machine learning on microcontrollers.

OpenEarable Dashboard. To make it easy to get started with recording sensor data, a dedicated dashboard for OpenEarable was developed (see Figure 6.4). Users can connect to the device via their browser, configure sampling rates, enable sensor streams and record as well as export sensor data as CSV files.

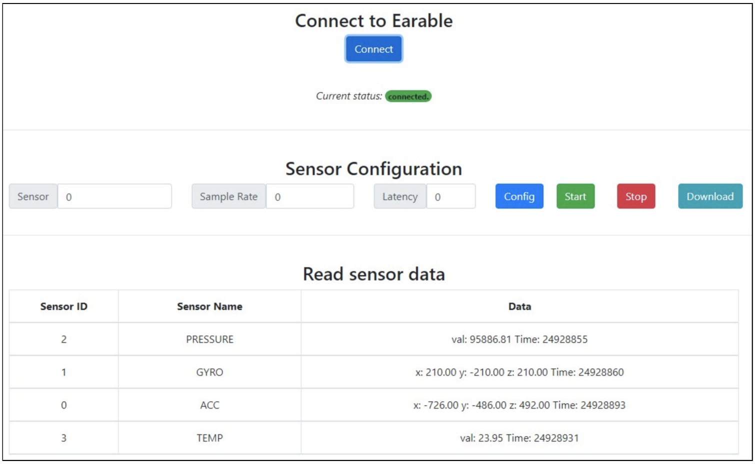


Figure 6.4.: OpenEarable dashboard that lets users configure sampling rates, enable sensors, and record data via WebBLE.

edge-ml. Out of the box, the OpenEarable firmware supports edge-ml[4], which is an open-source and browser-based toolchain for machine learning on microcontrollers. It offers recording, dataset management and labeling features. Using the default firmware installed on OpenEarable, users can simply connect to the device via WebBLE in their browser via edge-ml. In addition to data collection and labeling, it is also possible to train, validate, and export embedded machine learning models for OpenEarable using the edge-ml toolchain.

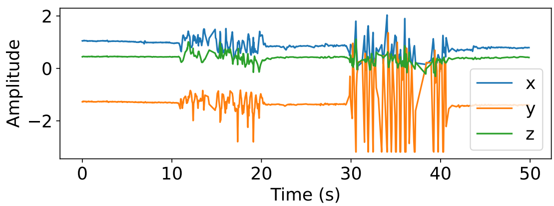
[4] edge-ml - `https://edge-ml.org`

6.1.4 Application Examples

To gain an understanding that the OpenEarable platform is outputting valid data, three example application scenarios from the literature were investigated.

6.1.4.1 Motion Tracking. Measuring motion on the ear is a common application in the earable space which can be used for a number of applications [455]. Figure 6.6 shows accelerometer and gyroscope readings from a single subject performing a sequence of three activities: (1) standing still, (2) walking, and (3) jumping jacks. These activities were chosen to elicit distinct patterns, and the jumping jacks allow us to validate the mechanical stability of the OpenEarable during vigorous motion. The activities were performed for 10 seconds in the following sequential order: stand, walk, stand, jumping jacks, and stand.

(a) Data obtained with the integrated accelerometer.



(b) Data obtained with the integrated gyroscope.

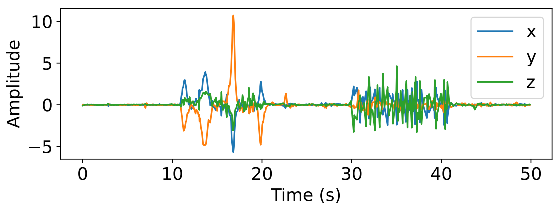
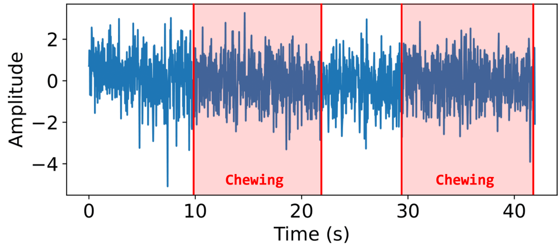


Figure 6.5.: Motion activities recorded with OpenEarable. The following events can be seen in chronological order: Standing still, walking, standing still, jumping jacks and standing still. The data was z-normalized before plotting.

6.1.4.2 Ear Canal Pressure. A popular ear canal pressure application is the detection of jaw motions [17]. Figure 6.6a shows two sequences of chewing activities, with a break in-between. The importance of an air tight ear canal can be seen with the chewing events clearly visible when using the sealed ear buds. The distinct pressure signal demonstrates the feasibility of in-ear barometry using the OpenEarable platform. In addition, this also shows how OpenEarable may be used for tensor tympani muscle interaction, as introduced in section 5.2 of this dissertation, which uses the same sensor (Bosch BMP280) for measuring ear canal pressure changes.

(a) No sealing is applied to the earpiece of OpenEarable (using standard silicone tips) which results in just noise being measured at the ear.



(b) Sealing is applied to the earpiece of OpenEarable using that expanding foam eartip so that the pressure sensor is air tight with the ear canal.

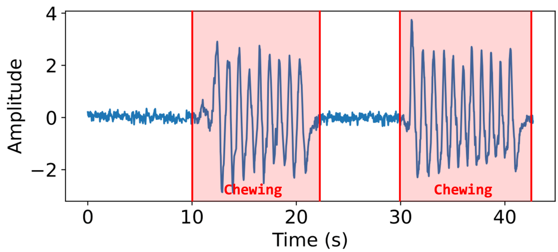


Figure 6.6.: A sequence of ear canal pressure changes including chewing and not chewing with (a) a standard conical silicone eartip, and (b) *Etymotic Research* disposable eartip ER1-14A. The data was z-normalized before plotting.

6.1.4.3 Ear Canal Sound Reflections. It is possible for the ultrasonic microphone to pick up an inaudible signal from the speaker. This information can be used to understand the shape of the ear canal because the sound is reflected differently depending on the shape, a principle which can be used for authentication [484]. While more detailed evaluations are necessary to assess generalized authentication performance based on OpenEarable, Figure 6.7 shows the spectogram of a 18kHz tone played in the ear canal for 8 seconds.

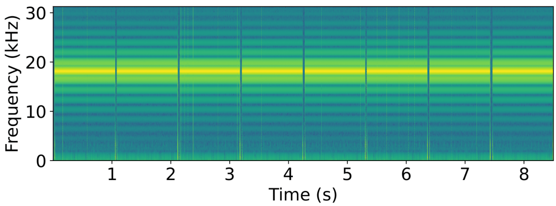


Figure 6.7.: Spectrogram of a reflected ultrasonic signal which was emitted into the ear canal with multiple 1s long probings.

6.1.4.4 Future Work. As this is the first version of OpenEarable, there are a number of limitations with the prototype and improvements to be made. Firstly, The earable can not be paired with a second device. So far, there are no libraries available for recording data from the OpenEarable platform on either Android or iOS devices which is a high priority item considering popular use cases of earable devices. While transferring a continuous audio signal over BLE is technically feasible, it is not yet implemented in the current OpenEarable firmware and the speaker only supports playback of a constant frequency. The Bluetooth classic advanced audio distribution profile (A2DP) is not yet supported. The standard ArduinoBLE library with current configuration achieved a transmission rates of 6.5 kB/s for the audio signal, but it is intended to use the NimBLE library, however, just recently support for the nRF52840 was added and compatibility with the bootloader of OpenEarable is pending but under active development. OpenEarable does not support reading out the battery level

6.1.5 Conclusion

OpenEarable is the first-if-its-kind open hardware initiative for earable research. Through this dissertation, a new device that features a series of sensors and actuators was introduced: a 9-axis accelerometer, gyroscope and magnetometer, an ear canal pressure and temperature sensor, an inward facing ultrasonic microphone as well as a speaker, a push button, and a controllable LED. The dissertation has shown the validity of the hardware based on three example application scenarios.

6.2 Designing and Prototyping

The following two sections study earables from the end user's perspective. Subsection 6.2.1 delves into how a series of commodity earphones are perceived in terms of wearability and comfort when used during sleep, which is a continuation of the possibility to wear earables for respiration rate tracking in the context of sleep apnea (see section 3.2, subsubsection 3.1.5.3, and subsubsection 3.1.2.2). Subsection 6.2.2 then takes the developers perspective to understand how earable prototyping platforms, such as OpenEarable (see section 6.1) or eSense [236], foster creativity and support problem-based learning.

6.2.1 Wearability and Comfort of Earables during Sleep

Many publications looked into sleep-related parameters such as sleep stages based on earable sensing (see subsubsection 3.1.2.2). However, none questioned their comfort and wearability during sleep. Therefore, a study was conducted to understand the wearability of 7 daytime targeted earables with 14 participants wearing each device for one night. All devices reduced subjective sleep quality and affected the sleeping position of the wearer. The following sections introduce related work, the study procedure, and results in depth, as published in ISWC 2021 [454].

6.2.1.1 Background and Related Work. Hoelzemann et al. [196] explored the wearability of the eSense earables during daytime activities. Overall, their comfort was well perceived. They also found that users were "generally not concerned about their appearance". Haas et al. [170] found that, e.g., in-ears are more suitable than over-ears and vice versa in some situations for personal soundscape curation. Xu et al. [495] evaluated the social acceptability of hand-based input on the ear, which revealed that tapping gestures are more appropriate than sliding.

According to Knight and Baber [247], a wearable can be characterized based on the comfort rating scale (CRS) across six dimensions: emotions, attachment, harm, perceived change, movement, and anxiety. Each dimension represents a deviation from the norm (i.e., not wearing a device). The presented study applies the CRS to assess wearability and comfort.

6.2.1.2 Study Design.

Earables. Two researchers performed a Google search (keyword "Bluetooth earphones") looking for market-available earable form factors targeted at daytime usage. The presented study looks into their suitability during sleep, implying 24/7 usage. This yielded the following basic attachment principles: in-ear, fixed between tragus and anti-tragus, concha filling, hook around the ear, and hook around the head. The earables use in the study span this design space (see Figure 6.8) and have different materials (e.g., hard / rigid and soft / bendy). Headphones were excluded from the study.

Procedure. Participants were given in-person instructions about the study. They wear the devices at their home during sleep. Participants give informed consent, and they are handed over the earables, questionnaires, and written instructions. The order of devices per participant is counterbalanced according to Williams design [490]. They are asked to wear every device for one night. If they feel too uncomfortable (e.g., sleeping impossible), participants can take off the device. They fill in the questionnaire immediately after waking up. Per earable, the questionnaire covers the CRS with the addition of "during sleep" for each item [247], questions on the influence and suitability of the device during sleep, if they kept the earable in and how many times it fell out, general concerns, demographics, and free text feedback. After completing all nights, participants rank all devices according to their preference.

Figure 6.8.: Earables selected for the study, ranging from in-ear to ear-hooks.

6.2.1.3 Results. For the study, 14 participants were recruited (8 male, 6 female, mean age: 27.0±3.1 years, sample of convenience). One participant was excluded because of incomplete replies. For significance testing, pairwise Wilcoxon signed-rank tests with Bonferroni correction for multiple comparisons were applied. Hypotheses were no differences of each device compared to all others per question item.

Comfort Rating Scale (CRS). Figure 6.9 summarizes the responses. Generally, users were not concerned about how they look during sleep, as to be expected by the pervasive nature of earables and the private setting. For attachment, Z8 was perceived more compared to K8 ($p < 0.1$). Interestingly, they were the heaviest (160 g) and lightest earable (2 x 8 g) of the selected devices. Regarding harm, in-ear devices that have a rigid part placed inside the concha (eSense, HolyHigh) were more painful to wear ($p < 0.1$) than in-ear devices that do not (Cosinuss). Intuitively, the perceived pain of wearing a particular earable and the number of participants that took the device out prematurely is closely coupled ($r_{xy} = 0.9$). Based on the participant feedback Z8 created discomfort at the temple. Devices wrapping around the ear which only have a soft or flexible or no part in the ear canal were perceived as less painful and taken out less often (see Table 6.2). Participants did not feel different or strange when wearing any of the selected earables. However, differences between the earables regarding the impairment on movement are well notable

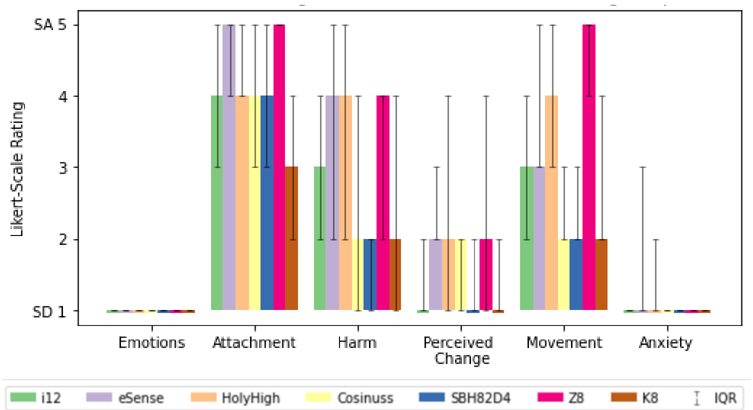


Figure 6.9.: Dimensions of the CRS for all earables [247]. SD = strongly disagree, SA = strongly agree. Results as median and inter-quartile range (IQR).

(see next paragraph). Overall, participants did not feel insecure wearing any of the earables.

Suitability for Sleep. Figure 6.9 (top) shows that all earables reduced the sleep quality of participants to some degree. However, SBH82D4 did less than i12, HolyHigh, and Z8 – and K8 less than Z8 ($p < 0.1$). This finding aligns with the advantages identified across the CRS dimensions. Regarding sleep position, it was found that SBH82D4 influences it less than eSense, i12, Z8, and HolyHigh ($p < 0.1$). In the context of wearables, acceptance plays an important role. Participants largely disagree with wearing any of the selected earables during sleep, see Figure 6.10 (top). However, they are less likely to reject SBH82D4 than eSense and HolyHigh, and Cosinuss, as well as K8, compared to eSense, i12, HolyHigh, and Z8 ($p < 0.1$). Zero up to six pointswere assigned to the participants' ranks at the end of the study. The summed up scores and respective ranks are presented in Table 6.2. The ranks align with what one would expect from the results presented before.

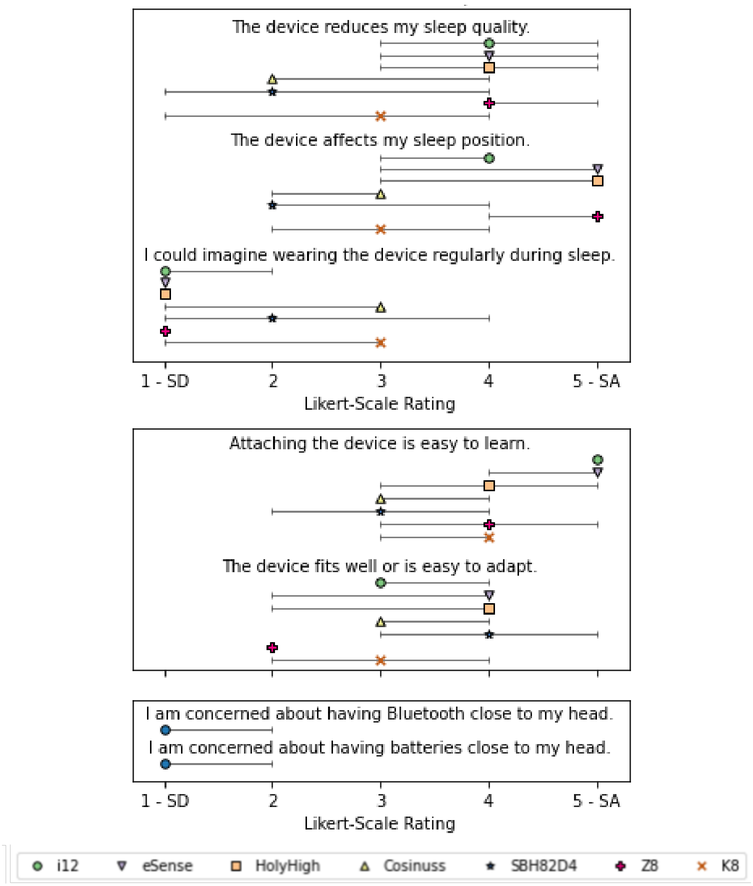


Figure 6.10.: Effects on sleep of the earables; (c) attachment of the earables and general concerns of the users. SD = strongly disagree, SA = strongly agree. Results as median and inter-quartile range (IQR).

Attachment and Concerns. Figure 6.10 (bottom) shows that all earables are easy to attach. eSense and i12 only attach in the ear canal and were easier to put on than SBH82D4, which uncommonly wraps under the ear ($p < 0.1$). Z8 fit worse than all other devices ($p < 0.1$) and participants were undecided whether Cosinuss, which can be fit to the ear, could be well adapted. Partic-

Table 6.2.: Ranking and times devices were taken out / fell off.

Device	Rank	Score	Fell Off	Taken Out
SBH82D4	1	64	6	4
K8	2	60	10	6
Cosinuss	3	41	2	6
i12	4	38	8	9
Z8	5	25	5	9
eSense	6	24	7	9
HolyHigh	7	21	5	10

ipants mainly were not concerned about having batteries close to their heads. Four users agreed with concerns about Bluetooth, but the majority still disagreed. Table 6.2 shows that especially the K8, i12, and eSense fell off during sleep – as to be expected given they are loosely attached.

Materials. The most favoured in-ear device (SBH82D4) is made from soft silicone material compared to, e.g., HolyHigh, which is primarily built from hard plastic. In contrast, the rigid K8 form factor was comfortable to support slim plastic parts behind the ear.

6.2.1.4 Conclusion. Overall, daytime-targeted earables applied in the study appeared unsuitable for usage during sleep as they reduced subjective sleep quality and affected the sleeping position. Primarily, rigid parts at the outer ear canal and the concha should be avoided. The area behind the ear offers real estate for rigid components, but the maximum weight is limited. Earables designed for sleep should be soft in front of the ear and in the ear canal (e.g., [161]) possibly combined with slim rigid components placed at the medial auricle (e.g., [375]). Users are not worried about how they look and do not feel insecure about earables during sleep. The results are limited by the earables that were used as they are not targeted at usage during sleep. This leaves opportunities for future work (e.g., *BOSE Sleepbuds*) with more participants over multiple nights per device to further validate and expand the presented findings.

6.2.2 Prototyping with Earables

In this section, the potential of earable prototyping platforms such as eSense [235] or OpenEarable (see section 6.1) is investigated. Making earabletechnology available to a broader audience might support innovation. A platform that aims to improve the availability of earable prototyping is *eSense* by Nokia Bell Labs [235]. Based on 24 publications, the design space of earable prototyping with eSense was characterized. Then, the *eSense* platform (6-axis IMU, auditory I/O) was evaluated for its problem-based learning usability with university students. Data was collected from 79 undergraduate students who developed 39 projects. Questionnaire-based results suggest that the platform creates interest in the subject matter and supports self-directed learning. The projects align with the research space, indicating ease of use, but lack contributions for more challenging topics. Additionally, many projects included games not present in current research. The average SUS score of the platform was 67.0. The majority of problems are technical issues (e.g., connecting, playing music). The following sections present the results as published in ISWC 2020 [451].

6.2.2.1 Background and Related Work. A platform already used for problem based education in various disciplines is *Arduino* [38]; however, building earables with Arduino requires substantial effort as wiring components, setting up connectivity interfaces, and fitting everything onto the ear is a challenging task already.

To summarize available earables and their data interfaces, *Google Search* was initially queried for “in-ear sensor” and “smart earbuds”. Relevant results were chosen from the first 20 entries (earbuds with more than audio I/O). The rationale for choosing *Search* as opposed to, e.g., *Scholar*, was that market available earables are new and heavily commercialized. Table 6.3 presents the results. Based on the earable name *Google Scholar* was queried to extract publications from the first 30 results (excluding patents / citations) that make use of any of the device capabilities.

Cosinuss° provides a highly engineered platform for medical applications with many sensors and multiple research publications [107; 403; 92; 220; 41; 393; 368]. However, data access requires proprietary software or unofficial, limited

Table 6.3.: Earable comparison. (✓≙ yes, (✓)≙ limited, ✗≙ no) (API ≙ open app framework, FW ≙ open firmware, MIC ≙ microphone, SPK ≙ speaker, IMU ≙ inertial measurement unit, PPG ≙ photoplythysmography, TMP ≙ temperature sensor, PRS ≙ pressure sensor)

Platform	API	FW	MIC	SPK	IMU	PPG	TMP	PRS
cosinuss°	(✓)	✗	✗	✗	✓	✓	✓	✗
Bose	✗	✓	✓	✓	✗	✓	✗	✗
Soul Blade	✗	✗	✓	✓	✓	✓	✗	✗
Jabra	✗	✗	✓	✓	✓	✓	✗	✗
Bragi	(✓)	(✓)	✗	✗	✗	✗	✗	✗
AirPods	✓	✗	✓	✓	(✓)	✗	✗	✗
eSense	✓	✗	✓	✓	✓	✗	✗	✗
OpenEarable	✓	✓	✓	✓	✓	(✓)	(✓)	✓

libraries. While the *Bose SoundSport* firmware is open, it lacks an application framework [64]. The *Jabra Elite Sports* have many sensors but no data access [208; 332], similar to *Blade* [127]. *Bragi* provides a powerful, closed developer framework but no more hardware [72]. Apple's AirPods have a proprietary API for iOS [20]. The Google Pixel or Amazon Echo Buds offer no access to the IMU. Platforms like eSense [235] and OpenEarable (see section 6.1), therefore, fill a critical gap for non-experts with a clearly defined, open framework with many research publications (see subsubsection 6.2.2.2). This simplification helps getting started with earables quickly.

A key motivation for making earables available to non-experts is to enable learning based on real-world projects. In that regard, the idea of problem-based learning (PBL) is to "*confront students with problems from practice which provides a stimulus for learning*" [68]. According to Munshi et al. [331], effective PBL should "*lead to thinking, analysis, and reasoning. It should stimulate self-directed learning and fit with students' prior knowledge. The problem should show clear links with the future profession and enhance interest in the subject matter.*". These underlying principles will guide the presented evaluation.

6.2.2.2 Design Space of Earable Prototyping. As shown in chapter 2, earables enable a multitude of compelling use cases. The following evaluation focuses on eSense, which at the time of conducting the presented study has been available for around two years. eSense is a stereo earable equipped with a microphone, 6-axis inertial measurement unit, and dual-mode Bluetooth. It allows recording of three real-time data streams - audio, motion, and proximity (BLE RSSI). It is powered by a CSR processor, has a 40 mAh battery, weighs 20 grams, and the dimension is 18 x 18 x 20 mm, including enclosure [235; 320]. Existing software libraries make it easy to get started with the earables. Using the paper by Kawsar et al. [235] and the *esense.io* website, seventeen peer-reviewed publications were identified that make direct use of the platform. In Figure 6.11, a fine granular research design space based on eSense publications is presented, with the primary motivation to use it for putting the non-expert projects into context that will be implemented by students. This design space differs from the presented high-level taxonomy in chapter 2 and presents only a subset of what is possible with earables in their entirety.

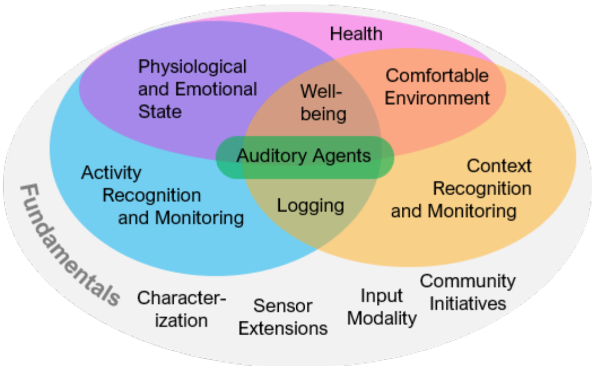


Figure 6.11.: Work areas of publications based on eSense to put the student projects into context.

In a *health* context, eSense was used to track the user's *physiological state* (respiration rate [448]). The platform was also used as mHealth building block [39] and to detect jaw clenching [405]. *Activity recognition* includes step counting [388], stay/walk detection as well as classifying speaking, eating and head shaking or nodding [202], drinking or chewing [282; 319], and exercising [206]. Additionally, frown and smile detection [265] as well as head movement [391] are possible indicators of the *emotional state*. Also, a data *logging* mechanism was proposed [196]. *Auditory agents* were used for auditive manipulation to support walking in a straight line [301]. Katayama et al. [230] proposed a setup for adapting a conversational agent's style, tone, and volume to the emotional, environmental, and social *context* to create a *comfortable environment*. *Well-being* creates a feedback loop, e.g., to understa conversational well-being [321]. In general, eSense can be an interface device of virtual conversational agents [2]. Work on *fundamental* principles creates the foundation of the research space. This includes e.g., *characterization* of wearing variability [323] and understanding earables as *input modality*, e.g., to control a robot arm [351]. *Community initiatives* include approaches for secure earable data sharing platforms [386], and also *sensor extensions* such as a magnetometer [134]. Publications for any of the other platforms also fall into the proposed prototyping design space.

6.2.2.3 Application Use Case Projects. To investigate problem based learning with eSense, undergrad university students were challenged to developed their own projects using the platform's sensing capabilities. They were asked to use the existing library [39] for *Flutter*, which is a cross-platform framework.

Course Setup and Data Collection. To collect data from students about their opinions and experience with the eSense prototyping platform, the course and data collection was structured as follows:

- all students fill in an initial questionnaire reporting their general demographics and attitude towards earables, based on related work on problem-based learning [331]. (i.e., if they see them as being relevant for their future profession)

- after filling in the questionnaire, the project task is presented and they are given a workshop on *Flutter* and the *eSense* library

- each student gets a pair of earbuds from the lab to take home and implements the project in a self-organized manner

- students finish the course by handing in their final project results and fill in a second questionnaire that asks about their experience (i.e., problems they faced and if the assignment helped to <3improve their knowledge in relevant areas)

Otherwise, students were not limited to any idea or particular challenge and asked them to build a creative solution that uses any of the eSense features.

6.2.2.4 Results. In total, 70 students participated in building projects. Out of all students, 63 were male and seven female. Sixty-eight participants were undergraduate computer science and two business informatics students. 22 took the university's HCI lecture. Twenty-four students never built a mobile app before, and only five used Flutter. Twelve have used Bluetooth before. Fifty-eight students reported never having heard of earables.

Suitability for Problem-based Learning. Figure 6.12 illustrates the results of four rating scales that students filled in before the project. Overall, there is no strong agreement if earables are relevant for the future profession or match the overall curriculum. Key factors might be different electives and individual career plans of students. Nevertheless, the majority of students agreed that the project was interesting.

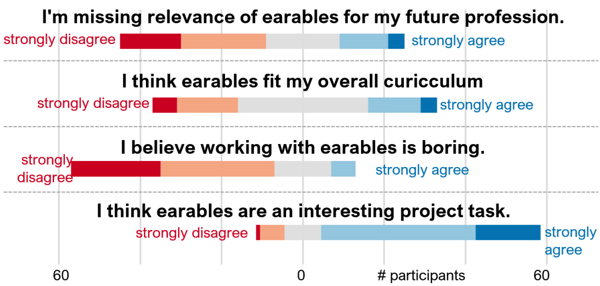


Figure 6.12.: Rating scales filled-in by students (N = 70) before implementing their projects with the eSense earables.

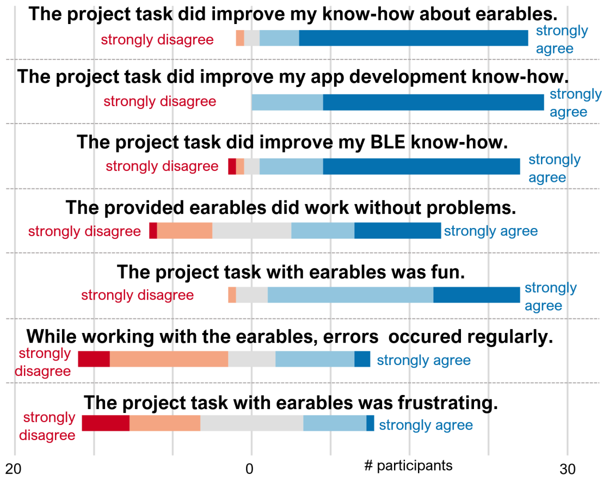


Figure 6.13.: Results of self-directed eSense projects (N = 37).

Retrospective Project Reflections. Unfortunately, due to COVID-19, only 37 students handed in the retrospective questionnaire. The results of the rating items are reported in Figure 6.13. Overall, the project has supported self-directed learning as students strongly agree that it improved their knowledge about earables, app development, and Bluetooth Low Energy. Even though the earables did not entirely work without problems and errors occurred regularly for some, the assignment was reported to have been fun. The task was not frustrating for the majority of students. THis shows how earables are an effective way to encourage problem-based learning.

Figure 6.14 reports the project types according to the categories defined in subsubsection 6.2.2.2. A topic not found in the design space analysis were games, for which earables served as direct input. Also, students built activity tracking (e.g., push-ups, situps, squats, walking, cycling, street-crossing, head-banging) with threshold state machines. One person classified nodding with ML. Auditory agents solely used text-to-speech. Some students implemented physiological state tracking (e.g., sleep position and work desk body pose). The auxiliary material contains detailed project descriptions.

Likely due to the higher complexity, no project included context recognition

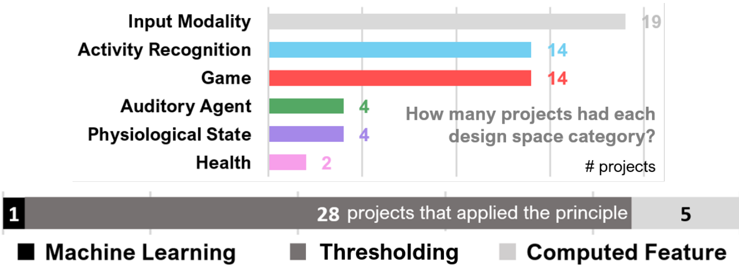


Figure 6.14.: Project results based on category and data processing approach (n = 37). Projects can have multiple categories.

(e.g., use proximity, or audio features).

The overall mean SUS score was 67.0 (max: 90.0, min: 35.0, sd: 14.2), which, according to literature, is close to the average reference value of 68 [413]. For good usability, however, the score should be above 80. Figure 6.15 the issues students faced during development extracted from free-text.

Connection issues included earables taking long to connect and problems connecting both earbuds. Seven users had problems with unstable audio. Issues with the button occurred when it involuntarily triggered music, started phone calls, or sent more than one press. Especially projects with more movement reported loose fit. Only three users reported trouble understanding sensor data. Issues with the library (iOS inconsistencies) and noisy data are also very little. In contrast to the technical problems, students expressed that “the earables look modern”, that it was “the best assignment they had at university so far”

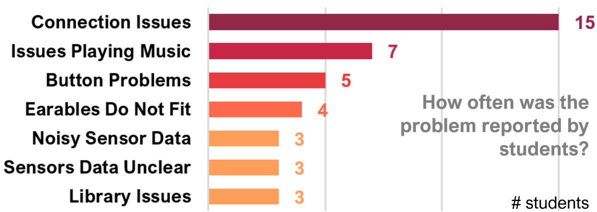


Figure 6.15.: Problems during development extracted from free text provided by participants (N = 37).

and also "an interesting introduction into pervasive computing". Overall, all students successfully fulfilled the assignment. Even though the complexity of results varies, eSense was easy to use by non-experts.

6.2.2.5 Conclusion. Overall, earables are enablers of a broad research design space. Undergraduate CS students are excited about the subject matter and enjoy as well as succeed in building applications with the prototyping platform eSense. However, relevance for their future career is limited. Frustration mostly roots in technical issues and little problems are reported with realizing ideas. The platform's documentation for teaching in a PBL setting seems sufficient. Considering the limited technical depth of student projects, teachers should encourage responsible design (i.e., in a health context). The reference material might also include core sensor concepts, and classification libraries could simplify prototyping.

7. Discussion

The following sections first discuss the state of earables in the four key domains: physiological parameters and health (section 7.1), movement and activity (section 7.2), interaction (section 7.3), and authentication and identification (section 7.4). Section 7.5 then goes on to discuss future opportunities and challenges of earables. Lastly, section 7.6 discusses the overarching aim of this dissertation: to break new ground through novel sensing capabilities and steer earables towards a general-purpose wearable sensing platform.

7.1 Physiological Parameters and Health

As summarized in this dissertation based on systematic literature review, earables can sense a multitude of body functions from 9 (out of 11) major systems of the human body, including skeletal (e.g., gait [32]), muscular (e.g., facial expressions [304]), nervous (e.g., brain activity [115]), endocrine (e.g., emotions [35]), cardiovascular (e.g., blood pressure [76]), respiratory (e.g., breathing see section 3.2), reproductive (e.g., ovulation [284]), immune (e.g., coughing in section 3.3). and digestive (e.g., food intake [150]) systems.

This dissertation succeeded in expanding upon the existing body of research by introducing the required processing chain to measure respiration rate and detect cough based on inertial sensors in the ear canal (see section 3.2 and section 3.3). While the performance for respiration rate tracking is limited when the user is moving, the results are promising for scenarios where the user is at rest. Similarly, predicting cough from inertial sensors has its limitations under motion-induced stress. However, the dissertation showed that, in principle, monitoring patients over longer duration still allows predicting if a person is in a state of increased cough frequency with a relatively weak classifier. In the future, combining multiple sensors placed on the ears may further boost the performance of respiration rate sensing [53] and cough prediction [341].

Already today, ear-based temperature monitoring is available in consumer products [107]. Similarly, heart rate sensing has found its way into the hands of consumers with commodity earables supporting the possibility to measure heart rate inside the ear canal [429; 107]. Still, the ear holds a lot more potential for many health-related parameters to be sensed which will need further validation and increased robustness.

In sum, earables are already showing promising results in the health domain, and as the sensing and processing capabilities continue to advance, earables are likely to play a vital role in at-home monitoring and even possibly diagnosis.

7.2 Movement and Activity

The ears present a relatively stable anchoring point from which to sense a wide range of movements and activities, in turn enabling recognition of exercises, sports, and other daily activities. Despite earables not achieving the best performance when used in isolation, they can support and complement other devices in the wider wearable eco-system [433], and the everyday use of ear-based form factors make them easier to integrate into everyday life than many other devices – an important factor for successful habit formation [36]. Beyond just classification, quantification of sensed phenomena can also support users to self-regulate healthy behaviours in their everyday life, from monitoring their levels of physical activity to tracking what food and drink they consume.

To expand upon the status quo in activity sensing with earables, this dissertation introduced earEOG as a new, more comfortable modality to track eye movements based on electrodes placed around the ears (see section 4.2). While the results are promising, especially for horizontal eye movements, this research is still evolving. Some hurdles to overcome are, for example, that electrodes will have to be integrated more naturally into headphones, e.g., using cloth electrodes; or that the sensing principle will have to be more robust against motion artifacts.

The vast majority of movement and activity tracking research relies on motion sensors such as accelerometers and gyroscopes. Such sensors have already found their way into consumer products, for example Apple's Airpods even offering an API to access the raw motion data [20]. Nonetheless, apps do not make full use of these capabilities.

Therefore, combined with the physiological parameters and health data available on the platform, earables have the potential to be a powerful self-tracking tool in the quantified self movement [285].

7.3 Interaction

In the context of human-computer interaction (HCI), earables present an exciting opportunity for unique and novel interaction techniques given the rich and diverse sensing capabilities available on the earable platform. The ear itself is easily and comfortably reached by the hands [494; 241], while the distinctive surface area creates opportunities for targeted interactions [272]. The possibility to bend the ear and its unique shape as well as flat surface area around it also opens up an interesting interaction design space [494; 241]. Beyond the ear itself, the earable platform can also be used to detect other modalities for interaction, including head position [16], facial gestures [304], mouth movements [435], and eye gaze [374; 57].

In this dissertation, the physiological nature of the ear was harnessed for a new interaction paradigm. By using the tensor tympani, a small inner ear muscle is used as a discreet, eyes- and hands-free input (see section 5.2). While this method is unique, the main limitation is that only a subpart of the population has the ability to control the muscle which suggests future research that may explore if and how this ability might be learned.

From a market perpsective, interaction with earables still is limited to clicking and tapping. In addition, head position tracking is used passively for spatial audio [20]. A major hurdle for consumer adoption is the heavy lifting required for advanced processing which is challenging for low power ear-based devices. In addition, the experimental techniques in human-computer interaction lack proven reliability and robustness in field settings which hinders adoption further.

Despite the current limitations in interaction techniques and market adoption, ongoing research and advancements in HCI hold promise for expanding the potential of earables, enabling more diverse and reliable interactions in the future.

7.4 Authentication and Identification

Mobile devices typically secure sensitive data using biometrics. Hence, the fourth research area identified in the earables taxonomy in chapter 2 is "Authentication and Identification". While no new application-specific contributions have been made by the dissertation, this section briefly discusses related earable works in this area, whereas a detailed summary can be found in the related survey paper published as part of this dissertation [455].

Existing earable approaches commonly make use of unique individual features of the user's ear, skull, body, and even brain activity or body motion sensed by ear-located principles. More traditional passcode-based methods based on rhythmic patterns are also used. These systems function in two main ways: verification, where a user's claimed identity is confirmed or rejected, and identification, which seeks to match the user to a set of pre-enrolled identities [336].

Shape-based authentication and identification rely on the uniqueness of the individual's ear, skull, and body (see section A.15). Ear authentication, for instance, uses in-ear and over-ear devices to measure sound reflections from the static [151; 9; 288; 21] or dynamic ear canal [483] or the entire ear [9; 118]. The performance of these methods can be influenced by varying wearing conditions and environmental disturbances [151; 288; 21]. A standout method is one that uses the response of the cochlea's basilar membrane to sound stimuli, offering superior performance [276]. Skull and body-based authentication methods exploit the unique structures and tissues of these parts of the body. For example, methods have been developed that use sounds produced by the body, which are propagated through the structures of the skull [152], or vibrations from the mandible that are detected by an accelerometer [274]. Another method measures a leakage current propagated through the body to an earable when the user touches a metal-encased laptop, showing promising results [121].

Motion-based authentication and identification methods harness the individual patterns of a person's gait [135] or head movements [104]. Brain-based methods, on the other hand, capture EEG responses during specific tasks to create a unique biometric profile [112; 111; 336].

An alternative to these biometric approaches is passcode-based authentication, which involves rhythmic tapping on the earbud, measured by an accelerometer [54].

While earable authentication techniques are a growing area of interest for their potential to provide seamless and continuous authentication, their current performance does not quite match that of traditional smartphone biometrics such as fingerprint and face recognition [19].

Despite the current performance gap compared to traditional smartphone biometrics, the growing interest and advancements in earable authentication techniques pave the way for seamless and continuous authentication experiences in the future.

7.5 Future Opportunities and Challenges

As demonstrated through this dissertation and Figure 2.1, earables are a growing research area. Despite technological advances, the limited space of the earable form factor creates, and will remain, an engineering challenge as components need to be miniaturised (e.g., [362; 35; 66]). The need for miniaturisation and small form factor introduces a number of potential constraints on the available computational performance, storage, and power on the earable device itself. From the literature review and studies conducted in this dissertation, several overarching opportunities and challenges were identified that need to be overcome in order for earables to realise their full potential across the four research areas identified. The statistics presented in the following sections are based on the body of research described in the taxonomy chapter 2.

7.5.1 Self-contained Platform

Earables have the potential to be a self-contained, light-weight wearable platform that does not depend on other devices due to the vast array of sensing principles available on the platform and the audio, and potentially haptic, output capabilities. However, in part due to the currently available computing resources on-board, it is common that other devices are used for further processing of the earable sensor data. Out of the research papers that were examined for the taxonomy, only 12 (4%) reported processing of sensor data on the earable device itself, with 22 studies (8%) using a smartphone for processing. The vast majority of research papers used higher powered computing devices (e.g., laptops and desktop computers). While this does not necessarily mean the systems developed could not run on lower-powered devices, it highlights a lack of research into earables as a self-contained platform.

7.5.2 Power Consumption

In addition to maximising the computational resources on the platform, power sources are also required to fit into the earable form factor which makes power consumption an important consideration (e.g., [287; 279; 256]). This requirement can be a considerable barrier for applications requiring frequent computationally expensive operations (e.g., machine learning models) or large and/or continuous data throughput which can quickly drain the battery. However, because the earable form factor is commonly used in wearable, mobile contexts there is the potential for energy to be harvested on the device itself, with preliminary research showing the feasibility of harvesting energy from jaw movements [116; 66] and thermoelectric generators [3]. Energy harvesting technologies may be required to supplement earable's on-board power capacity and extend battery life as the platform matures.

7.5.3 Wearable and Smart Device Ecosystem

Earables are ideally placed to, and will likely become, part of the commodity wearable and smart device ecosystem alongside smartphones and smartwatches. Earables can, and often do, offload processing onto other devices which can be beneficial for many applications in which other devices would likely be present anyway. Offloading sensor data provides access to additional computing resources, helps to extend the battery life of the earables, and allows earables to benefit from other devices providing higher-level contextual information about the user and their location. Researchers have also started to explore how earables can work in synergy with other wearables and smart devices. This emerging research space has yielded promising results which have shown how combining earables with other devices results in improved performance compared with the devices in isolation (e.g., [310; 157; 433]). It is important that the research community explores and understands how earables best fit into and complement the current and future wearable and smart device ecosystems. It may also be noted that the "self-contained" and "device ecosystem" research directions are not mutually exclusive, and each will play an important role in unlocking the potential of the earable platform.

7.5.4 Integration of Multiple Sensors

Earables can come in many different form factors, from subtle in-ear hearing aids to over-the-ear headphones and mechanically anchored earphones for exercising. In this dissertation multiple sensing principles were discussed (see Table 2.3) that have been demonstrated on the earable platform. However, different sensors and their required methods of attachment present a challenge when considering what can be simultaneously sensed on the ear by one device. An earable that encompasses multiple sensing modalities will inherently be limited by what attachments are viable for a given combination of sensing principles. This introduces trade-offs between the sensing capabilities and method of attachment, and the summaries of related works have shown how some sensors have different performance characteristics that are dependent on the form factor (e.g., electrode placement for earEOG, see section 4.2).

To date, research into earables equipped with multiple sensing modalities or solutions that span different application domains is limited. However, commercial offerings such as the eSense[1] or more recently Cosinuss Two[2] provide platforms that remove the need for bespoke hardware to be created and may allow researchers to consider the relationship of multiple solutions on a single platform. With the introduction of a fully open platform for ear-based sensing in this dissertation (see section 6.1), further opportunities for novel multi-modal sensing applications are created.

[1] eSense earable computing platform: https://www.esense.io/
[2] Cosinuss: https://store.cosinuss.com/

7.5.5 Assessment Heterogeneity

The ability to understand and compare the performance of different sensing principles on the earable platform is crucial. However, both Masè et al. [298] and Ne et al. [340] noted in their hearable reviews that the literature suffers from a wide variety of protocols and measurements that made comparisons between studies difficult [298; 340]. This dissertation found this issue extends to the research papers reviewed in this work and is evidenced in the Appendices. For example, research into heart rate (section A.1) and blood oxygen saturation (section A.2) shows multiple performance metrics being used.

Not only does assessment heterogeneity make it difficult to compare sensors across the earable literature, but it also limits and obfuscates the understanding of earables compared to other wearable and smart devices (e.g., smartphones or smartwatches) and gold standards. In this dissertation, where possible this gap was closed by providing gold standard comparison points, as does Ne et al. [340] in their review. However, it should be of primary importance moving forward for researchers investigating new and novel sensing principles on the earable platform to make a concerted effort to establish standard protocols for specific sensing principles and phenomena.

7.5.6 Ecological Validity

On top of assessment heterogeneity, earable research to date generally suffers from a lack of ecological validity. Work showing preliminary results and proofs of concepts is common, with over half (52%) of the papers surveyed having a study with less than 10 participants, and 48 (18%) papers' evaluations are based on a single user only. While preliminary research is important for the laying the foundations and exploring what is capable on the platform, there are significant questions being raised as to whether these innovations, often studied in limited lab conditions, will hold up in-the-wild during the mobile use cases and scenarios that earables will likely be used in.

Further analysis found that only 26 (< 10%) papers included field experiments in their evaluation. Researchers need to take into account the context of use which is likely going to be varied, dynamic, and possibly on-the-move in noisy environments or for prolonged periods. Researchers should consider any discomfort created by exerting force on the ear canal (e.g., [78; 290; 454]) or through blocking the hearing abilities of the user (e.g., [105]). Similarly, the form factor of the earable will affect the stability during different activities, but there also exists variability between wearing sessions (e.g., [151; 322]) and earables are susceptible to motion artefacts (e.g., [281; 57; 450; 184]), audio noise (e.g., [11; 487; 287]), and environmental weather conditions (e.g., [61]). Filling in these research gaps and embracing the opportunities they present will support and accelerate the transition from experimental prototypes to impactful, real-world products.

7.5.7 User Variability

Another issue related to ecological validity, which is also highlighted across the contributions in this dissertation, concerns user variability. Ergonomics research has shown that anthropometric differences exist between different genders and populations, including differences in upper ear height, concha width, lower ear height, and ear protrusion [267; 37; 147]. Another source of variability between users involves varying ear canal conditions which includes earwax blockages, ear infections, or other skin conditions that could affect the ear canal. These may affect some sensors more than others, for example those that require line-of-sight. Similarly, it is common for people to have ear piercings that may affect attachment. This is in addition to the potential problems introduced by attaching and maintaining contact with the sensors. Users may not be able to attach the sensors correctly first time (e.g., [49; 257; 322]) and there is the possibility that skin contact of a sensor is lost during use (e.g., [205]). This highlights the need for diverse research participation, and to understand how these factors affect wearability over long periods, and in the use contexts in which they were intended.

7.6 Aim and Objectives

This dissertation ambitiously strived to establish a foundational understanding of the capabilities of earables and to pave the way for versatile, general-purpose earable sensing. As such, earables present a unique opportunity to leverage a platform already embedded in our everyday lives through ear-based devices such as earphones and hearing aids. As a result, earables are socially acceptable to wear [46] and have the potential to be unobtrusive (e.g., [304; 264; 112]), discreet (e.g., [264; 279; 233]), inconspicuous (e.g., [160; 195; 105]), concealed (e.g., [115; 151]), privacy preserving (e.g., [264; 381; 261]), and non-stigmatizing [381; 240].

Building upon these fundamental properties, the primary objective of this dissertation was to comprehensively understand the existing capabilities of earable sensing technologies through a systematic literature review. This objective was accomplished through a uniquely comprehensive approach, reviewing over 270 relevant publications. However, it is important to note that earables are a rapidly progressing and broad field of research, which means that the developed taxonomy will continue to evolve as earables mature.

To expand on the existing research, the second objective was to identify novel sensing opportunities and explore under-explored phenomena. Specifically, the dissertation opportunistically investigated respiration rate sensing, cough detection, and earEOG based on available sensing principles and the global situation, particularly the COVID-19 pandemic. Furthermore, the author of this dissertation coincidentally possessed the necessary rumbling ability to utilize the novel interaction technique introduced, leading to further investigation. While these phenomena offer unique insights and advantages, the industry's willingness to incorporate such capabilities into commercial earables remains an open question for the future.

The third objective aimed to design and implement hardware and algorithms to detect these phenomena. Developing earables from scratch required deep technical knowledge and expertise. The dissertation successfully built relevant hardware and proposed suitable algorithms to detect the newly identified earable sensing capabilities. However, it is important to note that the devices and software are primarily intended for lab usage and research purposes. With the introduction of OpenEarable, the dissertation takes a significant step forward in bridging the gap between experimental research results and the sensing capabilities of commercially available ear-based devices.

To evaluate the reliability and usability of the proposed solutions, this dissertation extensively employed empirical studies. This ensured that the new earable devices and algorithms were tested to assess their generalizability and real-world applicability.

In conclusion, the objectives set by this dissertation were ambitious and not without significant challenges. However, tackling these challenges resulted in an enhanced understanding of earable sensing capabilities and catalyzed the development of general-purpose earable sensing.

8. Conclusion

The dissertation's main objective was to create fundamentally new insights into the capabilities of earables and guide them towards a general-purpose platform. This objective was achieved by breaking new ground and creating deeper understanding across four major domains: (i) physiological monitoring and health, (ii) movement and activity, (iii) interaction, and (iv) authentication and identification.

Evidently shown by this dissertation through systematic literature review, earables have broad sensing capabilities. Arguably, these capabilities are more diverse than any other location across the human body which makes earables unique and superior in comparison to any other commodity wearable platform. While the existing body of research is immense, this dissertation could still significantly expand upon existing knowledge. As such, the dissertation has provided significant contributions to the field, offering previously unseen sensing capabilities to detect novel phenomena: respiration, cough, eye gaze, and tensor tympani muscle contraction. From a methodological point of view, the presented work succeeded in identifying suitable sensing principles, collecting relevant data, and building methods to detect and quantify phenomena. On top, the dissertation accumulates in the OpenEarable platform which integrates diverse sensing capabilities to establish a general-purpose platform for the earable research field as a whole.

The detection of new phenomena gives rise to promising health-related respiration tracking and cough detection. The data processing techniques proposed to detect these physiological phenomena are simple enough to run continuously on earables which suggests that the generated insights can have a real-world impact. Hence, this dissertation could extend upon the ability of earables to be a platform that guides healthier lifestyles.

Additionally, monitoring eye gaze with electrodes placed around the ears opens up the opportunity for an entirely new, more comfortable and easily accessible eye tracking modality with much potential left to be explored. Integrating such sensing capabilities into commodity devices will allow deeper insights into human behavior through activity tracking and supports the vision of earables becoming a platform that integrates deeply with user behavior.

A novel interaction paradigm based on the voluntary contraction of the musculus tensor tympani in the human middle ear results in unique advantages. Compared to existing interaction techniques, it is entirely invisible to external observers and does not require the use of the eyes or hands. Through this contribution, the dissertation effectively transforms the ear from a purely receptive organ into a dual function one, enabling it not only to receive, but also to generate output signals. The principle is so special that users found it "magical and almost telepathic" which suggests that the dissertation unleashed significant unused potential of earables.

Lastly, the introduction of the open hardware OpenEarable platform reduces entry barriers into ear-based sensing research and bridges the gap between foundational research on earable sensing capabilities and their practical deployment in real-world scenarios.

While the findings are significant, their transition to consumer devices is still in the future. Nonetheless, some research is already finding its way into the real-world with other universities and major industry players showing interest in the presented research and collaborating to continue the journey that this dissertation embarks on.

In conclusion, the presented work marks a significant advancement in the field of earable computing and beyond. It provides crucial insights and sets the stage for further research. The future of earables is promising and the vision of earables as a general-purpose platform for human augmentation becomes increasingly apparent.

Bibliography

[1] H. Abdelnasser, K. A. Harras, and M. Youssef. UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 277–286. ACM, 2015.

[2] U. G. Acer, M. Van den Broeck, and F. Kawsar. The city as a personal assistant. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 1102–1106, 2019.

[3] R. Ahmed, M. U. Mehmood, A. Mannan, J. Y. Lee, S. H. Lim, W. Chun, et al. Design and Performance Analysis of Powering a Wireless Earphone by a Thermoelectric Generator. *IEEE Access*, 9:54457–54465, 2021.

[4] T. Ahmed, M. M. Rahman, M. Y. Ahmed, E. Nemati, M. Dinh, N. Folkman, J. Kuang, and A. Gao. RRMonitor: A Resource-Aware End-to-End System for Continuous Monitoring of Respiration Rate Using Earbuds. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 2463–2467. IEEE, 2021.

[5] J. Ahn, Y. Ku, D. Kim, J. Sohn, J.-H. Kim, and H. Kim. Wearable in-the-ear EEG system for SSVEP-based brain–computer interface. *Electronics Letters*, 54(7):413–414, 2018. doi: https://doi.org/10.1049/el.2017.3970. URL `https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/el.2017.3970`.

[6] A. Ahuja, A. Ferlini, and C. Mascolo. PilotEar: Enabling In-ear Inertial Navigation. In *Adjunct Proceedings of the 2021 ACM International*

Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers, pages 139–145, 2021.

[7] B. E. Ainsworth, W. L. Haskell, M. C. Whitt, M. L. Irwin, A. M. Swartz, S. J. Strath, W. L. O Brien, D. R. Bassett, K. H. Schmitz, P. O. Emplaincourt, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and science in sports and exercise*, 32(9; SUPP/1):S498–S504, 2000.

[8] S. Akiyama, K. Sato, Y. Makino, and T. Maeno. ThermOn: Thermo-Musical Interface for an Enhanced Emotional Experience. In *Proceedings of the 2013 International Symposium on Wearable Computers*, ISWC '13, page 45–52, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450321273. doi: 10.1145/2493988.2494326. URL `https://doi.org/10.1145/2493988.2494326`.

[9] A. H. M. Akkermans, T. A. M. Kevenaar, and D. W. E. Schobben. Acoustic ear recognition for person identification. In *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)*, pages 219–223, Oct 2005. doi: 10.1109/AUTOID.2005.11.

[10] Y. D. Alqurashi, T. Nakamura, V. Goverdovsky, J. Moss, M. I. Polkey, D. P. Mandic, and M. J. Morrell. A novel in-ear sensor to determine sleep latency during the Multiple Sleep Latency Test in healthy adults with and without sleep restriction. *Nature and science of sleep*, 10:385, 2018.

[11] T. Amesaka, H. Watanabe, and M. Sugimoto. Facial Expression Recognition Using Ear Canal Transfer Function. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC '19, page 1–9, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368704. doi: 10.1145/3341163.3347747. URL `https://doi.org/10.1145/3341163.3347747`.

[12] T. Amesaka, H. Watanabe, and M. Sugimoto. Facial expression recognition using ear canal transfer function. In *Proceedings of the 23rd International Symposium on Wearable Computers*, pages 1–9, 2019.

[13] O. Amft. A wearable earpad sensor for chewing monitoring. In *SENSORS, 2010 IEEE*, pages 222–227, Nov 2010. doi: 10.1109/ICSENS.2010.5690449.

[14] O. Amft, M. Kusserow, and G. Troster. Bite Weight Prediction From Acoustic Recognition of Chewing. *IEEE Transactions on Biomedical Engineering*, 56(6):1663–1672, June 2009. ISSN 1558-2531. doi: 10.1109/TBME.2009.2015873.

[15] D. R. Anderson, R. W. Whitmer, R. Z. Goetzel, R. J. Ozminkowski, J. Wasserman, S. Serxner, and H. E. R. O. H. R. Committee. The relationship between modifiable health risks and group-level health care expenditures. *American Journal of Health Promotion*, 15(1):45–52, 2000.

[16] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi. CanalSense: Face-Related Movement Recognition System Based on Sensing Air Pressure in Ear Canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, page 679–689, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450349819. doi: 10.1145/3126594.3126649. URL `https://doi.org/10.1145/3126594.3126649`.

[17] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi. Canalsense: Face-related movement recognition system based on sensing air pressure in ear canals. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pages 679–689, 2017.

[18] R. Angeli, M. Lise, C. Tabajara, and T. Maffacioli. Voluntary contraction of the tensor tympani muscle and its audiometric effects. *The Journal of Laryngology and Otology*, 127(12):1235, 2013.

[19] Apple. About touch ID advanced security technology, Sep 2017. URL `https://support.apple.com/en-us/HT204587`.

[20] Apple. AirPods Technical Description. https://www.apple.com/airpods-2nd-generation/specs/, 2020. Accessed: 2020-04-11.

[21] T. Arakawa, T. Koshinaka, S. Yano, H. Irisawa, R. Miyahara, and H. Imaoka. Fast and accurate personal authentication using ear acoustics. In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pages 1–4, Dec 2016. doi: 10.1109/APSIPA.2016.7820886.

[22] D. Ashbrook, C. Tejada, D. Mehta, A. Jiminez, G. Muralitharam, S. Gajendra, and R. Tallents. Bitey: An Exploration of Tooth Click Gestures for Hands-Free User Interface Control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '16, page 158–169, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450344081. doi: 10.1145/2935334.2935389. URL https://doi.org/10.1145/2935334.2935389.

[23] D. Ashbrook, C. Tejada, D. Mehta, A. Jiminez, G. Muralitharam, S. Gajendra, and R. Tallents. Bitey: An exploration of tooth click gestures for hands-free user interface control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 158–169, 2016.

[24] C. T. Association. Physical Activity Monitoring for Heart Rate ANSI/CTA-2065, 2018. URL https://standards.cta.tech/kwspub/published_docs/CTA-2065-Preview.pdf. [2018-12-02].

[25] L. Atallah, O. Aziz, B. Lo, and G. Yang. Detecting Walking Gait Impairment with an Ear-worn Sensor. In *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, pages 175–180, June 2009. doi: 10.1109/BSN.2009.41.

[26] L. Atallah, B. Lo, R. Ali, R. King, and G. Yang. Real-Time Activity Classification Using Ambient and Wearable Sensors. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):1031–1039, Nov 2009. ISSN 1558-0032. doi: 10.1109/TITB.2009.2028575.

[27] L. Atallah, G. G. Jones, R. Ali, J. J. H. Leong, B. Lo, and G. Yang. Observing Recovery from Knee-Replacement Surgery by Using Wearable Sensors. In *2011 International Conference on Body Sensor Networks*, pages 29–34, May 2011. doi: 10.1109/BSN.2011.10.

[28] L. Atallah, J. Leong, B. Lo, and G.-Z. Yang. Energy expenditure prediction using a miniaturized ear-worn sensor. *Medicine and science in sports and exercise*, 43(7):1369–1377, 2011.

[29] L. Atallah, B. Lo, R. King, and G. Yang. Sensor Positioning for Activity Recognition Using Wearable Accelerometers. *IEEE Transactions on Biomedical Circuits and Systems*, 5(4):320–329, Aug 2011. ISSN 1940-9990. doi: 10.1109/TBCAS.2011.2160540.

[30] L. Atallah, A. Wiik, G. G. Jones, B. Lo, J. P. Cobb, A. Amis, and G.-Z. Yang. Validation of an ear-worn sensor for gait monitoring using a force-plate instrumented treadmill. *Gait & posture*, 35(4):674–676, 2012.

[31] L. Atallah, O. Aziz, E. Gray, B. Lo, and G.-Z. Yang. An ear-worn sensor for the detection of gait impairment after abdominal surgery. *Surgical innovation*, 20(1):86–94, 2013.

[32] L. Atallah, A. Wiik, B. Lo, J. Cobb, A. Amis, and G. Yang. Gait asymmetry detection in older adults using a light ear-worn sensor. *Physiological measurement*, 35(5):N29, 2014.

[33] L. Atallah, C. Ciuhu, C. Wang, E. Bongers, T. Blom, I. Paulussen, and G. Noordergraaf. An ergonomic wearable core body temperature sensor. In *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 70–73, March 2018. doi: 10.1109/BSN.2018.8329661.

[34] C. Athavipach, S. Pan-ngum, and P. Israsena. Development of Low-Cost in-the-Ear EEG Prototype. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 1–6, July 2018. doi: 10.1109/JCSSE.2018.8457324.

[35] C. Athavipach, S. Pan-ngum, and P. Israsena. A Wearable In-Ear EEG Device for Emotion Monitoring. *Sensors*, 19(18), 2019. ISSN 1424-8220. doi: 10.3390/s19184014. URL `https://www.mdpi.com/1424-8220/19/18/4014`.

[36] C. Attig and T. Franke. Abandonment of personal quantification: A review and empirical study investigating reasons for wearable activity tracking attrition. *Computers in Human Behavior*, 102:223–237, 2020.

[37] K. Ban and E. S. Jung. Ear shape categorization for ergonomic product design. *International Journal of Industrial Ergonomics*, page 102962, 2020.

[38] M. Banzi and M. Shiloh. *Getting started with Arduino: the open source electronics prototyping platform*. Maker Media, Inc., 2014.

[39] J. E. Bardram. The CAMS eSense Framework: Enabling Earable Computing for mHealth Apps and Digital Phenotyping. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 3–7, 2019.

[40] P. Barralon, I. R. M. Martins, N. Merkle, and S. Schwarz. Augmented Hearing Assistance for Elderly People. 2015.

[41] P. Barralon, I. R. M. Martins, N. Merkle, and S. Schwarz. Augmented Hearing Assistance for elderly people. 2015.

[42] M. L. Barron and R. J. Fehring. Basal body temperature assessment: is it useful to couples seeking pregnancy? *MCN: The American Journal of Maternal/Child Nursing*, 30(5):290–296, 2005.

[43] D. R. Bassett, L. P. Toth, S. R. LaMunion, and S. E. Crouter. Step counting: a review of measurement considerations and health-related applications. *Sports Medicine*, 47(7):1303–1315, 2017.

[44] A. Bates, M. J. Ling, J. Mann, and D. Arvind. Respiratory rate and flow waveform estimation from tri-axial accelerometer data. In *2010 International Conference on Body Sensor Networks*, pages 144–150. IEEE, 2010.

[45] C. Bech Christensen, R. K. Hietkamp, J. M. Harte, T. Lunner, and P. Kidmose. Toward EEG-assisted hearing aids: Objective threshold estimation based on ear-EEG in subjects with sensorineural hearing loss. *Trends in hearing*, 22:2331216518816203, 2018.

[46] A. Bedri, D. Byrd, P. Presti, H. Sahni, Z. Gue, and T. Starner. Stick It in Your Ear: Building an in-Ear Jaw Movement Sensor. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, UbiComp/ISWC'15 Adjunct, page 1333–1338, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450335751. doi: 10.1145/2800835.2807933. URL `https://doi.org/10.1145/2800835.2807933`.

[47] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner. A Wearable System for Detecting Eating Activities with Proximity Sensors in the Outer Ear. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC '15, page 91–92, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450335782. doi: 10.1145/2802083.2808411. URL `https://doi.org/10.1145/2802083.2808411`.

[48] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner. Detecting Mastication: A Wearable Approach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ICMI '15, page 247–250, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450339124. doi: 10.1145/2818346.2820767. URL `https://doi.org/10.1145/2818346.2820767`.

[49] A. Bedri, R. Li, M. Haynes, R. P. Kosaraju, I. Grover, T. Prioleau, M. Y. Beh, M. Goel, T. Starner, and G. Abowd. EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), September 2017. doi: 10.1145/3130902. URL `https://doi.org/10.1145/3130902`.

[50] V. Bellotti, M. Back, W. K. Edwards, R. E. Grinter, A. Henderson, and C. Lopes. Making sense of sensing systems: five questions for designers and researchers. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 415–422, 2002.

[51] J. Bergström and K. Hornbæk. Human–Computer Interaction on the Skin. *ACM Comput. Surv.*, 52(4), August 2019. ISSN 0360-0300. doi: 10.1145/3332166. URL `https://doi.org/10.1145/3332166`.

[52] A. R. Bertelsen, H. Bladt, C. B. Christensen, S. L. Kappel, H. O. Toft, M. L. Rank, K. B. Mikkelsen, and P. Kidmose. Generic Dry-Contact Ear-EEG. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5552–5555, July 2019. doi: 10.1109/EMBC.2019.8857351.

[53] A. Bestbier and P. R. Fourie. Development of a vital signs monitoring wireless ear probe. In *2018 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC)*, pages 1–5, April 2018. doi: 10.1109/SAIBMEC.2018.8363196.

[54] H. Bi, Y. Sun, J. Liu, and L. Cao. SmartEar: Rhythm-Based Tap Authentication Using Earphone in Information-Centric Wireless Sensor Network. *IEEE Internet of Things Journal*, 9(2):885–896, 2021.

[55] S. Bi, T. Wang, E. Davenport, R. Peterson, R. Halter, J. Sorber, and D. Kotz. Toward a Wearable Sensor for Eating Detection. In *Proceedings of the 2017 Workshop on Wearable Systems and Applications*, WearSys '17, page 17–22, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450349598. doi: 10.1145/3089351.3089355. URL `https://doi.org/10.1145/3089351.3089355`.

[56] S. Bi, T. Wang, N. Tobias, J. Nordrum, S. Wang, G. Halvorsen, S. Sen, R. Peterson, K. Odame, K. Caine, R. Halter, J. Sorber, and D. Kotz. Auracle: Detecting Eating Episodes with an Ear-Mounted Sensor. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(3), September 2018. doi: 10.1145/3264902. URL `https://doi.org/10.1145/3264902`.

[57] M. G. Bleichner and S. Debener. Concealed, unobtrusive ear-centered EEG acquisition: cEEGrids for transparent EEG. *Frontiers in human neuroscience*, 11:163, 2017.

[58] M. G. Bleichner, M. Lundbeck, M. Selisky, F. Minow, M. Jäger, R. Emkes, S. Debener, and M. De Vos. Exploring miniaturized EEG electrodes for brain-computer interfaces. An EEG you do not see? *Physiological Reports*, 3(4):e12362, 2015. doi: https://doi.org/10.14814/phy2.12362. URL `https://physoc.onlinelibrary.wiley.com/doi/abs/10.14814/phy2.12362`.

[59] M. G. Bleichner, B. Mirkovic, and S. Debener. Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison. *Journal of neural engineering*, 13(6):066004, 2016.

[60] K. Blinowska and P. Durka. Electroencephalography (eeg). *Wiley encyclopedia of biomedical engineering*, 2006.

[61] C. A. Boano, M. Lasagni, and K. Römer. Non-invasive measurement of core body temperature in Marathon runners. In *2013 IEEE International Conference on Body Sensor Networks*, pages 1–6, May 2013. doi: 10.1109/BSN.2013.6575484.

[62] M. Bock, U. Hohlfeld, K. Von Engeln, P. A. Meier, J. Motsch, and A. J. Tasman. The accuracy of a new infrared ear thermometer in patients undergoing cardiac surgery. *Canadian Journal of Anesthesia*, 52(10): 1083–1087, 2005.

[63] E. Bondareva, E. R. Hauksdóttir, and C. Mascolo. Earables for Detection of Bruxism: a Feasibility Study. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 146–151, 2021.

[64] Bose. Bose SoundSport Firmware. `https://github.com/bosefirmware/ced`, 2020. Accessed: 2020-04-11.

[65] L. Bouarfa, L. Atallah, R. M. Kwasnicki, C. Pettitt, G. Frost, and G. Yang. Predicting Free-Living Energy Expenditure Using a Miniaturized Ear-Worn Sensor: An Evaluation Against Doubly Labeled Water. *IEEE Transactions on Biomedical Engineering*, 61(2):566–575, Feb 2014. ISSN 1558-2531. doi: 10.1109/TBME.2013.2284069.

[66] J. Bouchard-Roy, A. Delnavaz, and J. Voix. In-Ear Energy Harvesting: Evaluation of the Power Capability of the Temporomandibular Joint. *IEEE Sensors Journal*, 20(12):6338–6345, June 2020. ISSN 1558-1748. doi: 10.1109/JSEN.2020.2976925.

[67] W. Boucsein. *Electrodermal activity*. Springer Science & Business Media, 2012.

[68] D. Boud and G. Feletti. *The challenge of problem-based learning*. Psychology Press, 1998.

[69] A. Boukhayma, A. Barison, S. Haddad, and A. Caizzone. Earbud-embedded micro-power mm-sized optical sensor for accurate heart beat monitoring. *IEEE Sensors Journal*, 21(18):19967–19977, 2021.

[70] R. E. Bouserhal, T. H. Falk, and J. Voix. Improving the quality of in-ear microphone speech via adaptive filtering and artificial bandwidth extension. *MANUSCRIPT-BASED THESIS PRESENTED TO*, page 25, 2016.

[71] B. Bradke and B. Everman. Investigation of Photoplethysmography Behind the Ear for Pulse Oximetry in Hypoxic Conditions with a Novel Device (SPYDR). *Biosensors*, 10(4):34, 2020.

[72] Bragi. Bragi Software. `https://www.bragi.com`, 2020. Accessed: 2020-04-11.

[73] L. Britt, A. B. Peitzman, P. S. Barie, and G. J. Jurkovich. *Acute care surgery*. Springer, 2012.

[74] J. D. Bronzino and D. R. Peterson. *Biomedical engineering fundamentals*. CRC press, 2014.

[75] F. Brudy, C. Holz, R. Rädle, C.-J. Wu, S. Houben, C. N. Klokmose, and N. Marquardt. Cross-device taxonomy: survey, opportunities and challenges of interactions spanning across multiple devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–28, 2019.

[76] N. Bui, N. Pham, J. J. Barnitz, Z. Zou, P. Nguyen, H. Truong, T. Kim, N. Farrow, A. Nguyen, J. Xiao, R. Deterding, T. Dinh, and T. Vu. EBP: A Wearable System For Frequent and Comfortable Blood Pressure Monitoring From User's Ear. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450361699. doi: 10.1145/3300061.3345454. URL `https://doi.org/10.1145/3300061.3345454`.

[77] N. Bui, N. Pham, J. J. Barnitz, Z. Zou, P. Nguyen, H. Truong, T. Kim, N. Farrow, A. Nguyen, J. Xiao, et al. ebp: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. In *The 25th annual international conference on mobile computing and networking*, pages 1–17, 2019.

[78] V. Buil and G. Hollemans. Acceptable operating force for buttons on in-ear type headphones. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pages 186–189, Oct 2005. doi: 10.1109/ISWC.2005.7.

[79] V. Buil, G. Hollemans, and S. van de Wijdeven. Headphones with Touch Control. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices amp; Services*, MobileHCI '05, page 377–378, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 1595930892. doi: 10.1145/1085777.1085877. URL `https://doi.org/10.1145/1085777.1085877`.

[80] V. Buil, G. Hollemans, and S. van de Wijdeven. Headphones with touch control. In *Proceedings of the 7th international conference on Human*

computer interaction with mobile devices & services, pages 377–378, 2005.

[81] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster. Robust recognition of reading activity in transit using wearable electrooculography. In *Pervasive Computing: 6th International Conference, Pervasive 2008 Sydney, Australia, May 19-22, 2008 Proceedings 6*, pages 19–37. Springer, 2008.

[82] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster. Eye movement analysis for activity recognition. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 41–50, 2009.

[83] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster. Eye movement analysis for activity recognition using electrooculography. *IEEE transactions on pattern analysis and machine intelligence*, 33(4):741–753, 2010.

[84] C. P. Burgos, L. Gärtner, M. A. G. Ballester, J. Noailly, F. Stöcker, M. Schönfelder, T. Adams, and S. Tassani. In-Ear Accelerometer-Based Sensor for Gait Classification. *IEEE Sensors Journal*, 20(21):12895–12902, Nov 2020. ISSN 1558-1748. doi: 10.1109/JSEN.2020.3002589.

[85] G. Cao, K. Yuan, J. Xiong, P. Yang, Y. Yan, H. Zhou, and X.-Y. Li. Earphonetrack: involving earphones into the ecosystem of acoustic motion tracking. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 95–108, 2020.

[86] Y. Cao, H. Chen, F. Li, and Y. Wang. CanalScan: Tongue-Jaw Movement Recognition via Ear Canal Deformation Sensing. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*, pages 1–10. IEEE, 2021.

[87] S. K. Card, T. P. Moran, and A. Newell. The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7):396–410, 1980.

[88] J. Carioli, A. Delnavaz, R. J. Zednik, and J. Voix. Piezoelectric Earcanal Bending Sensor. *IEEE Sensors Journal*, 18(5):2060–2067, March 2018. ISSN 1558-1748. doi: 10.1109/JSEN.2017.2783299.

[89] N. Celik, W. Balachandran, N. Manivannan, E. Winter, B. Schnalzer, and H. Burgsteiner. Wearable mobile ear-based ECG monitoring device using graphene-coated sensors. In *2017 IEEE SENSORS*, pages 1–3, Oct 2017. doi: 10.1109/ICSENS.2017.8233911.

[90] P. Celka, C. Verjus, R. Vetter, P. Renevey, and V. Neuman. Motion resistant earphone located infrared based heart rate measurement device. *Biomedical Engineering*, 2004.

[91] J. S. Chaglla E., N. Celik, and W. Balachandran. Measurement of Core Body Temperature Using Graphene-Inked infrared Thermopile Sensor. *Sensors*, 18(10), 2018. ISSN 1424-8220. doi: 10.3390/s18103315. URL `https://www.mdpi.com/1424-8220/18/10/3315`.

[92] J. S. Chaglla E, N. Celik, W. Balachandran, et al. Measurement of core body temperature using graphene-inked infrared thermopile sensor. *Sensors*, 18(10):3315, 2018.

[93] L. Chan, R.-H. Liang, M.-C. Tsai, K.-Y. Cheng, C.-H. Su, M. Y. Chen, W.-H. Cheng, and B.-Y. Chen. FingerPad: Private and Subtle Interaction Using Fingertips. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, page 255–260, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450322683. doi: 10.1145/2501988.2502016. URL `https://doi.org/10.1145/2501988.2502016`.

[94] I. Chatterjee, M. Kim, V. Jayaram, S. Gollakota, I. Kemelmacher, S. Patel, and S. M. Seitz. ClearBuds: wireless binaural earbuds for learning-based speech enhancement. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, pages 384–396, 2022.

[95] T. Chen and C. Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939785. URL `http://doi.acm.org/10.1145/2939672.2939785`.

[96] T. Chen, B. Steeper, K. Alsheikh, S. Tao, F. Guimbretière, and C. Zhang. 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*, pages 112–125, 2020.

[97] V. Chen, X. Xu, R. Li, Y. Shi, S. Patel, and Y. Wang. Understanding the Design Space of Mouth Microgestures. In *Designing Interactive Systems Conference 2021*, pages 1068–1081, 2021.

[98] Y.-C. Chen, C.-Y. Liao, S.-w. Hsu, D.-Y. Huang, and B.-Y. Chen. Exploring User Defined Gestures for Ear-Based Interactions. *Proceedings of the ACM on Human-Computer Interaction*, 4(ISS):1–20, 2020.

[99] H. Chennamma and X. Yuan. A survey on eye-gaze tracking techniques. *arXiv preprint arXiv:1312.6410*, 2013.

[100] I. Choi and D. Kim. Toning: New Experience of Sharing Music Preference with Interactive Earphone in Public Space. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '20, page 533–538, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450361071. doi: 10.1145/3374920.3374983. URL `https://doi.org/10.1145/3374920.3374983`.

[101] I. Choi, S. Rajaram, L. A. Varghese, and B. G. Shinn-Cunningham. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Frontiers in human neuroscience*, 7:115, 2013.

[102] R. R. Choudhury. Earable computing: A new area to think about. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, pages 147–153, 2021.

[103] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr. Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package). *Neurocomputing*, 307:72–77, 2018.

[104] C. Clarke, P. Ehrich, and H. Gellersen. Motion Coupling of Earable Devices in Camera View. In *19th International Conference on Mobile and Ubiquitous Multimedia*, pages 13–17, 2020.

[105] Cong-Zhi Wang and Yong-Ping Zheng. Home-Telecare of the elderly living alone using an new designed ear-wearable sensor. In *2008 5th International Summer School and Symposium on Medical Devices and Biosensors*, pages 71–74, June 2008. doi: 10.1109/ISSMDBS.2008.4575019.

[106] S. Cooper. Relationship of hearing protector type and prevalence of external auditory canal pathology. In *Am. Ind. Hyg. Asso. Conf. Las Vegas, NV, paper*, volume 23, 1985.

[107] Cosinuss. Coninuss Earables. `https://www.cosinuss.com/products/`, 2020. Accessed: 2020-04-11.

[108] E. Costanza, S. A. Inverso, and R. Allen. Toward subtle intimate interfaces for mobile devices using an EMG controller. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 481–489, 2005.

[109] M. A. Cretikos, R. Bellomo, K. Hillman, J. Chen, S. Finfer, and A. Flabouris. Respiratory rate: the neglected vital sign. *Medical Journal of Australia*, 188(11):657–659, 2008.

[110] T. K. Cureton Jr. Bodily posture as an indicator of fitness. *Research Quarterly. American Association for Health, Physical Education and Recreation*, 12(sup2):348–367, 1941.

[111] M. T. Curran, J. Yang, N. Merrill, and J. Chuang. Passthoughts authentication with low cost EarEEG. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1979–1982, Aug 2016. doi: 10.1109/EMBC.2016.7591112.

[112] M. T. Curran, N. Merrill, J. Chuang, and S. Gandhi. One-Step, Three-Factor Authentication in a Single Earpiece. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, UbiComp '17, page 21–24, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450351904. doi: 10.1145/3123024.3123087. URL `https://doi.org/10.1145/3123024.3123087`.

[113] G. de Graaf, D. Kuratomi Cruz, J. C. Haartsen, F. Hooijschuur, and P. J. French. Heart Rate Extraction in a Headphone Using Infrared Thermometry. *IEEE Transactions on Biomedical Circuits and Systems*, 13(5):1052–1062, Oct 2019. ISSN 1940-9990. doi: 10.1109/TBCAS.2019.2930312.

[114] A. De Luca, R. Weiss, and H. Drewes. Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN-Entry. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces*, OZCHI '07, page 199–202, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595938725. doi: 10.1145/1324892.1324932. URL `https://doi.org/10.1145/1324892.1324932`.

[115] S. Debener, R. Emkes, M. De Vos, and M. Bleichner. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Scientific reports*, 5(1):1–11, 2015.

[116] A. Delnavaz and J. Voix. Piezo-earpiece for micro-power generation from ear canal dynamic motion. *Journal of Micromechanics and microengineering*, 23(11):114001, 2013.

[117] A. Dementyev and J. A. Paradiso. WristFlex: low-power gesture input with wrist-worn pressure sensors. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 161–166, 2014.

[118] M. Derawi. Biometric acoustic ear recognition. In *2016 International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–4, Dec 2016. doi: 10.1109/BIOSMART.2016.7835597.

[119] M. Dhuliawala, J. Lee, J. Shimizu, A. Bulling, K. Kunze, T. Starner, and W. Woo. Smooth eye movement interaction using EOG glasses. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pages 307–311, 2016.

[120] N. K. Dim and X. Ren. Investigation of suitable body parts for wearable vibration feedback in walking navigation. *International Journal of Human-Computer Studies*, 97:34–44, 2017. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2016.08.002. URL `https://www.sciencedirect.com/science/article/pii/S107158191630088X`.

[121] D. Ding, L. Yang, Y.-C. Chen, and G. Xue. Leakage or Identification: Behavior-irrelevant User Identification Leveraging Leakage Current on Laptops. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–23, 2021.

[122] G. Djupesland. Middle ear muscle reflexes elicited by acoustic and nonacoustic stimulation. *Acta oto-laryngologica*, 57(sup188):287–292, 1964.

[123] K. Doddabasappla and R. Vyas. Statistical and Machine-learning based Recognition of Coughing Events using Tri-axial Accelerometer Sensor data from Multiple Wearable points. *IEEE Sensors Letters*, 2021.

[124] K. Doddabasappla and R. Vyas. Spectral summation with machine learning analysis of tri-axial acceleration from multiple wearable points

on human body for better cough detection. *IEEE Sensors Letters*, 5(9): 1–4, 2021.

[125] H. Dong, P. M. Matthews, and Y. Guo. A new soft material based in-the-ear EEG recording technique. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5709–5712. IEEE, 2016.

[126] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, C. Valderrama, P. Lebecque, and T. Dutoit. Objective Study of Sensor Relevance for Automatic Cough Detection. *IEEE Journal of Biomedical and Health Informatics*, 17(3):699–707, 2013.

[127] S. Electronics. BLADE True Wireless Earphones. `https://www.soulelectronics.com/products/blade-sports-ai-voice-coach-true-wireless`, 2020. Accessed: 2020-04-11.

[128] X. Fan, L. Shangguan, S. Rupavatharam, Y. Zhang, J. Xiong, Y. Ma, and R. E. Howard. HeadFi: bringing intelligence to all headphones. In *MobiCom*, pages 147–159, 2021.

[129] L. Fang, **T. Röddiger**, F. Schmid, and M. Beigl. EarRecorder: A Multi-Device Earable Data Collection Toolkit. In *Augmented Humans Conference 2021*, pages 286–288, 2021.

[130] A. Favre-Félix, C. Graversen, T. Dau, and T. Lunner. Real-time estimation of eye gaze by in-ear electrodes. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4086–4089. IEEE, 2017.

[131] A. Favre-Felix, R. Hietkamp, C. Graversen, T. Dau, and T. Lunner. Steering of audio input in hearing aids by eye gaze through electrooculography. In *Proceedings of the International Symposium on Auditory and Audiological Research*, volume 6, pages 135–142, 2017.

[132] A. Favre-Félix, C. Graversen, T. A. Bhuiyan, M. A. Skoglund, S. Rotger-Griful, M. L. Rank, T. Dau, and T. Lunner. Absolute eye gaze estimation with biosensors in hearing aids. *Frontiers in neuroscience*, 13:1294, 2019.

[133] A. Ferlini, A. Montanari, C. Mascolo, and R. Harle. Head Motion Tracking Through In-Ear Wearables. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 8–13, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369022. doi: 10.1145/3345615.3361131. URL `https://doi.org/10.1145/3345615.3361131`.

[134] A. Ferlini, A. Montanari, C. Mascolo, and R. Harle. Head Motion Tracking Through in-Ear Wearables. 2019.

[135] A. Ferlini, D. Ma, R. Harle, and C. Mascolo. EarGate: gait-based user identification with in-ear microphones. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 337–349, 2021.

[136] A. Ferlini, A. Montanari, A. Grammenos, R. Harle, and C. Mascolo. Enabling In-Ear Magnetic Sensing: Automatic and User Transparent Magnetometer Calibration. In *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–8. IEEE, 2021.

[137] A. Ferlini, A. Montanari, C. Min, H. Li, U. Sassi, and F. Kawsar. In-Ear PPG for Vital Signs. *IEEE Pervasive Computing*, pages 1–10, 2021. doi: 10.1109/MPRV.2021.3121171.

[138] F. F. Ferri. *Ferri's Clinical Advisor 2010 E-Book: 5 Books in 1*. Elsevier Health Sciences, 2009.

[139] L. Fiedler, J. Obleser, T. Lunner, and C. Graversen. Ear-EEG allows extraction of neural responses in challenging listening scenarios — A future technology for hearing aids? In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology*

Society (EMBC), pages 5697–5700, Aug 2016. doi: 10.1109/EMBC.2016.7592020.

[140] J. M. Fontana, M. Farooq, and E. Sazonov. Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior. *IEEE Transactions on Biomedical Engineering*, 61(6):1772–1779, June 2014. ISSN 1558-2531. doi: 10.1109/TBME.2014.2306773.

[141] C. for Devices, U. F. Radiological Health, and D. Administration. Pulse oximeter accuracy and limitations, 2016. URL `https://www.fda.gov/medical-devices/safety-communications/pulse-oximeter-accuracy-and-limitations-fda-safety-communication`.

[142] E. S. Ford, T. J. Cunningham, and J. B. Croft. Trends in self-reported sleep duration among US adults from 1985 to 2012. *Sleep*, 38(5):829–832, 2015.

[143] C. Frauenberger, T. Stockman, and M.-L. Bourguet. A survey on common practice in designing audio in the user interface. In *Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster, UK 21*, pages 1–9, 2007.

[144] K. N. Frayn. *Metabolic regulation: a human perspective*. John Wiley & Sons, 2009.

[145] E. Freeman, G. Griffiths, and S. A. Brewster. Rhythmic Micro-Gestures: Discreet Interaction on-the-Go. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, page 115–119, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450355438. doi: 10.1145/3136755.3136815. URL `https://doi.org/10.1145/3136755.3136815`.

[146] S. A. L. Frohn, J. S. Matharu, and J. A. Ward. Towards a Characterisation of Emotional Intent during Scripted Scenes Using In-Ear Movement Sensors. In *Proceedings of the 2020 International Symposium on Wearable Computers*, ISWC '20, page 37–39, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450380775.

doi: 10.1145/3410531.3414292. URL `https://doi.org/10.1145/3410531.3414292`.

[147] F. Fu and Y. Luximon. A systematic review on ear anthropometry and its industrial design applications. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 30(3):176–194, 2020.

[148] P. Gallego Cascón, D. J. Matthies, S. Muthukumarana, and S. Nanayakkara. ChewIt. An Intraoral Interface for Discreet Interactions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. doi: 10.1145/3290605.3300556. URL `https://doi.org/10.1145/3290605.3300556`.

[149] H. Gamper, S. Tervo, and T. Lokki. Head orientation tracking using binaural headset microphones. In *Audio Engineering Society Convention 131*. Audio Engineering Society, 2011.

[150] Y. Gao, N. Zhang, H. Wang, X. Ding, X. Ye, G. Chen, and Y. Cao. iHear Food: Eating Detection Using Commodity Bluetooth Headsets. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 163–172, June 2016. doi: 10.1109/CHASE.2016.14.

[151] Y. Gao, W. Wang, V. V. Phoha, W. Sun, and Z. Jin. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(3), September 2019. doi: 10.1145/3351239. URL `https://doi.org/10.1145/3351239`.

[152] Y. Gao, Y. Jin, J. Chauhan, S. Choi, J. Li, and Z. Jin. Voice In Ear: Spoofing-Resistant and Passphrase-Independent Body Sound Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–25, 2021.

[153] S. Gashi, A. Saeed, A. Vicini, E. Di Lascio, and S. Santini. Hierarchical Classification and Transfer Learning to Recognize Head Gestures and

Facial Expressions Using Earbuds. In *Proceedings of the 2021 International Conference on Multimodal Interaction*, pages 168–176, 2021.

[154] L. Geddes, M. Voelz, C. Combs, D. Reiner, and C. F. Babbs. Characterization of the oscillometric method for measuring indirect blood pressure. *Annals of biomedical engineering*, 10(6):271–280, 1982.

[155] E. Ghomi, G. Faure, S. Huot, O. Chapuis, and M. Beaudouin-Lafon. Using rhythmic patterns as an input method. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1253–1262, 2012.

[156] A. Gibaldi, M. Vanegas, P. J. Bex, and G. Maiello. Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behavior research methods*, 49:923–946, 2017.

[157] J. Gong, X. Zhang, Y. Huang, J. Ren, and Y. Zhang. Robust Inertial Motion Tracking through Deep Sensor Fusion across Smart Earbuds and Smartphone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2):1–26, 2021.

[158] T. Govaert, G. Dinant, K. Aretz, and J. Knottnerus. The predictive value of influenza symptomatology in elderly people. *Family practice*, 15(1): 16–22, 1998.

[159] V. Goverdovsky, D. Looney, P. Kidmose, and D. P. Mandic. In-Ear EEG From Viscoelastic Generic Earpieces: Robust and Unobtrusive 24/7 Monitoring. *IEEE Sensors Journal*, 16(1):271–277, Jan 2016. ISSN 1558-1748. doi: 10.1109/JSEN.2015.2471183.

[160] V. Goverdovsky, W. Von Rosenberg, T. Nakamura, D. Looney, D. J. Sharp, C. Papavassiliou, M. J. Morrell, and D. P. Mandic. Hearables: Multimodal physiological in-ear sensing. *Scientific reports*, 7(1):1–10, 2017.

[161] V. Goverdovsky, W. Von Rosenberg, T. Nakamura, D. Looney, D. J. Sharp, C. Papavassiliou, M. J. Morrell, and D. P. Mandic. Hearables:

Multimodal physiological in-ear sensing. *Scientific reports*, 7(1):1–10, 2017.

[162] M. J. Grenness, J. Osborn, and W. L. Weller. Mapping ear canal movement using area-based surface matching. *The Journal of the Acoustical Society of America*, 111(2):960–971, 2002.

[163] T. Grosse-Puppendahl, C. Holz, G. Cohn, R. Wimmer, O. Bechtold, S. Hodges, M. S. Reynolds, and J. R. Smith. Finding common ground: A survey of capacitive sensing in human-computer interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3293–3315, 2017.

[164] G. E. Grossman, R. J. Leigh, L. A. Abel, D. J. Lanska, and S. Thurston. Frequency and velocity of rotational head perturbations during locomotion. *Experimental brain research*, 70(3):470–476, 1988.

[165] Y. Gu, E. Cleeren, J. Dan, K. Claes, W. Van Paesschen, S. Van Huffel, and B. Hunyadi. Comparison between Scalp EEG and Behind-the-Ear EEG for Development of a Wearable Seizure Detection System for Patients with Focal Epilepsy. *Sensors*, 18(1), 2018. ISSN 1424-8220. doi: 10.3390/s18010029. URL `https://www.mdpi.com/1424-8220/18/1/29`.

[166] M. Guermandi, S. Benatti, V. J. Kartsch Morinigo, and L. Bertini. A Wearable Device for Minimally-Invasive Behind-the-Ear EEG and Evoked Potentials. In *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, Oct 2018. doi: 10.1109/BIOCAS.2018.8584814.

[167] A. Guillot and V. Thorey. RobustSleepNet: Transfer learning for automated sleep staging at scale. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1441–1451, 2021.

[168] D. G. Guo, F. E. H. Tay, L. Xu, L. M. Yu, M. N. Nyan, F. W. Chong, K. L. Yap, and B. Xu. A Long-term Wearable Vital Signs Monitoring System using BSN. In *2008 11th EUROMICRO Conference on Digital*

System Design Architectures, Methods and Tools, pages 825–830, Sep. 2008. doi: 10.1109/DSD.2008.126.

[169] U. Ha, Y. Lee, H. Kim, T. Roh, J. Bae, C. Kim, and H. Yoo. A Wearable EEG-HEG-HRV Multimodal System With Simultaneous Monitoring of tES for Mental Health Management. *IEEE Transactions on Biomedical Circuits and Systems*, 9(6):758–766, Dec 2015. ISSN 1940-9990. doi: 10.1109/TBCAS.2015.2504959.

[170] G. Haas, E. Stemasov, and E. Rukzio. Can't You Hear Me? Investigating Personal Soundscape Curation. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, pages 59–69, 2018.

[171] M. Haescher, J. Trimpop, D. J. Matthies, G. Bieber, B. Urban, and T. Kirste. aHead: considering the head position in a multi-sensory setup of wearables to recognize everyday activities with intelligent sensor fusions. In *International Conference on Human-Computer Interaction*, pages 741–752. Springer, 2015.

[172] M. Haescher, D. J. Matthies, J. Trimpop, and B. Urban. SeismoTracker: upgrade any smart wearable to enable a sensing of heart rate, respiration rate, and microvibrations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2209–2216. ACM, 2016.

[173] G. Hammour, M. Yarici, W. v. Rosenberg, and D. P. Mandic. Hearables: Feasibility and Validation of In-Ear Electrocardiogram. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5777–5780, July 2019. doi: 10.1109/EMBC.2019.8857547.

[174] G. M. Hammour and D. P. Mandic. Hearables: Making Sense from Motion Artefacts in Ear-EEG for Real-Life Human Activity Classification. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 6889–6893. IEEE, 2021.

[175] M. Hansenne. The p300 cognitive event-related potential. II. Individual variability and clinical application in psychopathology. *Neurophysiologie clinique= Clinical neurophysiology*, 30(4):211–231, 2000.

[176] T. Hao, C. Bi, G. Xing, R. Chan, and L. Tu. Mindfulwatch: A smartwatch-based system for real-time respiration monitoring during meditation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):57, 2017.

[177] J. Harris, S. Vance, O. Fernandes, A. Parnandi, and R. Gutierrez-Osuna. Sonic respiration: controlling respiration rate through auditory biofeedback. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 2383–2388. ACM, 2014.

[178] D. Harrison, P. Marshall, N. Bianchi-Berthouze, and J. Bird. Activity tracking: barriers, workarounds and customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 617–621, 2015.

[179] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[180] J. Hartvigsen, C. Leboeuf-Yde, S. Lings, and E. H. Corder. Is sitting-while-at-work associated with low back pain? A systematic, critical literature review. *Scandinavian journal of public health*, 28(3):230–239, 2000.

[181] O. Hashem, K. Alkiek, M. Youssef, and K. A. Harras. Leveraging Earables for Natural Calibration-Free Multi-Device Identification in Smart Environments. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, pages 92–98, 2021.

[182] D. D. He. *A wearable heart monitor at the ear using ballistocardiogram (BCG) and electrocardiogram (ECG) with a nanowatt ECG heartbeat detection circuit*. PhD thesis, Massachusetts Institute of Technology, 2013.

[183] D. D. He, E. S. Winokur, T. Heldt, and C. G. Sodini. The ear as a location for wearable vital signs monitoring. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 6389–6392, Aug 2010. doi: 10.1109/IEMBS.2010.5627309.

[184] D. D. He, E. S. Winokur, and C. G. Sodini. An Ear-Worn Vital Signs Monitor. *IEEE Transactions on Biomedical Engineering*, 62(11):2547–2552, Nov 2015. ISSN 1558-2531. doi: 10.1109/TBME.2015.2459061.

[185] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, F. Sun, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*, 2020.

[186] S. Hermann, P. Breitling, **T. Röddiger**, and M. Beigl. Cardiopulmonary Resuscitation Support: Comparison of Wrist-, Chest-, and Ear-Worn Devices and Estimation Algorithms. In *2021 International Symposium on Wearable Computers*, pages 28–32, 2021.

[187] S. Hermann, P. Breitling, **T. Röddiger**, and M. Beigl. Cardiopulmonary Resuscitation Support: Comparison of Wrist-, Chest-, and Ear-Worn Devices and Estimation Algorithms. In *2021 International Symposium on Wearable Computers*, pages 28–32, 2021.

[188] J. Hernandez, Y. Li, J. M. Rehg, and R. W. Picard. Cardiac and Respiratory Parameter Estimation Using Head-mounted Motion-sensitive Sensors. *EAI Endorsed Trans. Pervasive Health Technol.*, 1(1):e2, 2015.

[189] J. Hernandez, D. McDuff, and R. W. Picard. Biowatch: estimation of heart and breathing rates from wrist motions. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pages 169–176. IEEE, 2015.

[190] H. Higashi, T. M. Rutkowski, Y. Washizawa, A. Cichocki, and T. Tanaka. EEG auditory steady state responses classification for the

novel BCI. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4576–4579. IEEE, 2011.

[191] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1063–1072, 2014.

[192] K. Hinckley and D. Wigdor. Input technologies and techniques. *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, pages 151–168, 2002.

[193] L. Hládek, B. Porr, and W. O. Brimijoin. Real-time estimation of horizontal gaze angle by saccade integration using in-ear electrooculography. *Plos one*, 13(1):e0190420, 2018.

[194] J. A. Hodgkinson, M.-M. Lee, S. Milner, P. Bradburn, R. Stevens, F. R. Hobbs, C. Koshiaris, S. Grant, J. Mant, and R. J. McManus. Accuracy of blood-pressure monitors owned by patients with hypertension (ACCURATE study): a cross-sectional, observational study in central England. *British Journal of General Practice*, 70(697):e548–e554, 2020. doi: 10.3399/bjgp20x710381.

[195] A. Hoelzemann, H. Odoemelem, and K. Van Laerhoven. Using an In-Ear Wearable to Annotate Activity Data across Multiple Inertial Sensors. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 14–19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369022. doi: 10.1145/3345615.3361136. URL `https://doi.org/10.1145/3345615.3361136`.

[196] A. Hoelzemann, H. Odoemelem, and K. Van Laerhoven. Using an in-ear wearable to annotate activity data across multiple inertial sensors. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 14–19, 2019.

[197] H.-E. Holst, S. Ingelstedt, and U. Örtegren. Ear drum movements following stimulation of the middle ear muscles. *Acta Oto-Laryngologica*, 56(sup182):73–83, 1963.

[198] D. Hong, B. Zhang, Q. Li, S. Nirjon, R. Dickerson, G. Shen, X. Jiang, and J. Stankovic. SEPTIMU: Continuous in-Situ Human Wellness Monitoring and Feedback Using Sensors Embedded in Earphones. In *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, IPSN '12, page 159–160, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450312271. doi: 10.1145/2185677.2185727. URL `https://doi.org/10.1145/2185677.2185727`.

[199] J. Y. Hong, J. He, B. Lam, R. Gupta, and W.-S. Gan. Spatial audio for soundscape design: recording and reproduction. *Applied sciences*, 7(6): 627, 2017.

[200] D. Hossain, M. H. Imtiaz, and E. Sazonov. Comparison of Wearable Sensors for Estimation of Chewing Strength. *IEEE Sensors Journal*, 20(10):5379–5388, May 2020. ISSN 1558-1748. doi: 10.1109/JSEN.2020.2968009.

[201] J. L. Hossain and C. M. Shapiro. The prevalence, cost implications, and management of sleep disorders: an overview. *Sleep and Breathing*, 6 (02):085–102, 2002.

[202] T. Hossain, M. S. Islam, M. A. R. Ahad, and S. Inoue. Human activity recognition using earable device. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 81–84, 2019.

[203] J. Hsu, R. Stone, R. Logan-Sinclair, M. Worsdell, C. Busst, and K. Chung. Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder. *European Respiratory Journal*, 7(7):1246–1253, 1994.

[204] D.-Y. Huang, T. Seyed, L. Li, J. Gong, Z. Yao, Y. Jiao, X. A. Chen, and X.-D. Yang. Orecchio: Extending Body-Language through Actuated Static and Dynamic Auricular Postures. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, page 697–710, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450359481. doi: 10.1145/3242587.3242629. URL `https://doi.org/10.1145/3242587.3242629`.

[205] M. Huang, T. Tamura, T. Yoshimura, T. Tsuchikawa, and S. Kanaya. Wearable deep body thermometers and their uses in continuous monitoring for daily healthcare. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 177–180, Aug 2016. doi: 10.1109/EMBC.2016.7590669.

[206] S. Ishii, K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez. ExerSense: Real-Tme Physical Exercise Segmentation, Classification, and Counting Algorithm Using an IMU Sensor. *arXiv preprint arXiv:2004.10026*, 2020.

[207] S. Ishimaru, K. Kunze, K. Kise, J. Weppner, A. Dengel, P. Lukowicz, and A. Bulling. In the blink of an eye: combining head motion and eye blink frequency for activity recognition with google glass. In *Proceedings of the 5th augmented human international conference*, pages 1–4, 2014.

[208] Jabra. Elite Sport. `https://www.jabra.com.de/sports-headphones/jabra-elite-sport`, 2020. Accessed: 2020-04-11.

[209] N. K. Jacob, E. Balaban, R. Saunders, J. C. Batchelor, S. G. Yeates, and A. J. Casson. An Exploration of Behind-the-Ear ECG Signals From a Single Ear Using Inkjet Printed Conformal Tattoo Electrodes. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1283–1286, July 2018. doi: 10.1109/EMBC.2018.8512538.

[210] D. Jarchi, B. Lo, E. Ieong, D. Nathwani, and G. Yang. Validation of the e-AR Sensor for Gait Event Detection Using the Parotec Foot Insole with Application to Post-Operative Recovery Monitoring. In *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*, pages 127–131, June 2014. doi: 10.1109/BSN.2014.16.

[211] D. Jarchi, C. Wong, R. M. Kwasnicki, B. Heller, G. A. Tew, and G. Yang. Gait Parameter Estimation From a Miniaturized Ear-Worn Sensor Using Singular Spectrum Analysis and Longest Common Subsequence. *IEEE Transactions on Biomedical Engineering*, 61(4):1261–1273, April 2014. ISSN 1558-2531. doi: 10.1109/TBME.2014.2299772.

[212] D. Jarchi, C. Wong, R. M. Kwasnicki, B. Heller, G. A. Tew, and G.-Z. Yang. Gait parameter estimation from a miniaturized ear-worn sensor using singular spectrum analysis and longest common subsequence. *IEEE Transactions on Biomedical Engineering*, 61(4):1261–1273, 2014.

[213] D. Jarchi, B. Lo, C. Wong, E. Ieong, D. Nathwani, and G. Yang. Gait Analysis From a Single Ear-Worn Sensor: Reliability and Clinical Evaluation for Orthopaedic Patients. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(8):882–892, Aug 2016. ISSN 1558-0210. doi: 10.1109/TNSRE.2015.2477720.

[214] D. Jeong, J. Jeong, Y. Chae, and H. Choi. Identification of Attention State for Menu-Selection using In-Ear EEG Recording. In *2017 5th International Winter Conference on Brain-Computer Interface (BCI)*, pages 112–114, Jan 2017. doi: 10.1109/IWW-BCI.2017.7858176.

[215] Y. Jin, Y. Gao, Y. Zhu, W. Wang, J. Li, S. Choi, Z. Li, J. Chauhan, A. K. Dey, and Z. Jin. SonicASL: An Acoustic-based Sign Language Gesture Recognizer Using Earphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2):1–30, 2021.

[216] L. Jing, Z. Cheng, Y. Zhou, J. Wang, and T. Huang. Magic Ring: a self-contained gesture input device on finger. In *Proceedings of the 12th*

International Conference on Mobile and Ubiquitous Multimedia, pages 1–4, 2013.

[217] N. Johanna, H. Citrawijaya, and G. Wangge. Mass screening vs lockdown vs combination of both to control COVID-19: A systematic review. *Journal of public health research*, 9(4):jphr–2020, 2020.

[218] H. G. Jones, N. T. Greene, and W. A. Ahroon. Human middle-ear muscles rarely contract in anticipation of acoustic impulses: Implications for hearing risk assessments. *Hearing research*, 378:53–62, 2019.

[219] J. Juez, D. Henao, F. Segura, R. Gómez, M. Le Van Quyen, and M. Valderrama. Development of a wearable system with In-Ear EEG electrodes for the monitoring of brain activities: An application to epilepsy. In *2021 IEEE 2nd International Congress of Biomedical Engineering and Bioengineering (CI-IB&BI)*, pages 1–4. IEEE, 2021.

[220] D. Junger, N. M. Madrid, N. P. Malek, and C. Thies. Open wearables mobile platform to support personalized medicine. *Telemedicine and Telemonitoring in AAL Home Environments*, page 7, 2019.

[221] V. Kakaraparthi, Q. Shao, C. J. Carver, T. Pham, N. Bui, P. Nguyen, X. Zhou, and T. Vu. FaceSense: Sensing Face Touch with an Ear-worn System. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(3):1–27, 2021.

[222] N. Kaongoen and S. Jo. An auditory P300-based brain-computer interface using Ear-EEG. In *2018 6th International Conference on Brain-Computer Interface (BCI)*, pages 1–4, Jan 2018. doi: 10.1109/IWW-BCI.2018.8311519.

[223] S. L. Kappel and P. Kidmose. Study of impedance spectra for dry and wet EarEEG electrodes. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3161–3164, Aug 2015. doi: 10.1109/EMBC.2015.7319063.

[224] S. L. Kappel and P. Kidmose. High-density ear-EEG. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2394–2397, July 2017. doi: 10.1109/EMBC.2017.8037338.

[225] S. L. Kappel and P. Kidmose. Real-Life Dry-Contact Ear-EEG. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5470–5474, July 2018. doi: 10.1109/EMBC.2018.8513532.

[226] S. L. Kappel, C. B. Christensen, K. B. Mikkelsen, and P. Kidmose. Reference configurations for ear-EEG steady-state responses. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5689–5692, Aug 2016. doi: 10.1109/EMBC.2016.7592018.

[227] S. L. Kappel, D. Looney, D. P. Mandic, and P. Kidmose. Physiological artifacts in scalp EEG and ear-EEG. *Biomedical engineering online*, 16 (1):1–16, 2017.

[228] S. L. Kappel, M. L. Rank, H. O. Toft, M. Andersen, and P. Kidmose. Dry-Contact Electrode Ear-EEG. *IEEE Transactions on Biomedical Engineering*, 66(1):150–158, Jan 2019. ISSN 1558-2531. doi: 10.1109/TBME.2018.2835778.

[229] S. Katayama, A. Mathur, M. van den Broeck, T. Okoshi, J. Nakazawa, and F. Kawsar. Situation-Aware Emotion Regulation of Conversational Agents with Kinetic Earables. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 725–731, Sep. 2019. doi: 10.1109/ACII.2019.8925449.

[230] S. Katayama, A. Mathur, M. Van den Broeck, T. Okoshi, J. Nakazawa, and F. Kawsar. Situation-Aware Emotion Regulation of Conversational Agents with Kinetic Earables. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 725–731. IEEE, 2019.

[231] C. Katsini, Y. Abdrabou, G. E. Raptis, M. Khamis, and F. Alt. The Role of Eye Gaze in Security and Privacy Applications: Survey and Future HCI Research Directions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–21, 2020.

[232] J. Kavanagh, R. Barrett, and S. Morrison. The role of the neck and trunk in facilitating head stability during walking. *Experimental brain research*, 172(4):454, 2006.

[233] R. Kaveh, J. Doong, A. Zhou, C. Schwendeman, K. Gopalan, F. Burghardt, A. C. Arias, M. Maharbiz, and R. Muller. A Wireless, Multielectrode, User-generic Ear EEG Recording System. In *2019 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, Oct 2019. doi: 10.1109/BIOCAS.2019.8918700.

[234] R. Kaveh, J. Doong, A. Zhou, C. Schwendeman, K. Gopalan, F. L. Burghardt, A. C. Arias, M. M. Maharbiz, and R. Muller. Wireless User-Generic Ear EEG. *IEEE Transactions on Biomedical Circuits and Systems*, 14(4):727–737, Aug 2020. ISSN 1940-9990. doi: 10.1109/TBCAS.2020.3001265.

[235] F. Kawsar, C. Min, A. Mathur, and A. Montanari. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, 17(3):83–89, 2018.

[236] F. Kawsar, C. Min, A. Mathur, A. Montanari, U. G. Acer, and M. Van den Broeck. eSense: Open Earable Platform for Human Sensing. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pages 371–372, 2018.

[237] P. Khanna, T. Srivastava, S. Pan, S. Jain, and P. Nguyen. JawSense: recognizing unvoiced sound using a low-cost ear-worn system. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, pages 44–49, 2021.

[238] P. Kidmose, D. Looney, and D. P. Mandic. Auditory evoked responses from Ear-EEG recordings. In *2012 Annual International Conference of*

the IEEE Engineering in Medicine and Biology Society, pages 586–589, Aug 2012. doi: 10.1109/EMBC.2012.6345999.

[239] P. Kidmose, D. Looney, L. Jochumsen, and D. P. Mandic. Ear-EEG from generic earpieces: A feasibility study. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 543–546, July 2013. doi: 10.1109/EMBC.2013.6609557.

[240] P. Kidmose, D. Looney, M. Ungstrup, M. L. Rank, and D. P. Mandic. A Study of Evoked Potentials From Ear-EEG. *IEEE Transactions on Biomedical Engineering*, 60(10):2824–2830, Oct 2013. ISSN 1558-2531. doi: 10.1109/TBME.2013.2264956.

[241] T. Kikuchi, Y. Sugiura, K. Masai, M. Sugimoto, and B. H. Thomas. EarTouch: Turning the Ear into an Input Surface. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450350754. doi: 10.1145/3098279.3098538. URL `https://doi.org/10.1145/3098279.3098538`.

[242] T. Kikuchi, Y. Sugiura, K. Masai, M. Sugimoto, and B. H. Thomas. EarTouch: turning the ear into an input surface. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–6, 2017.

[243] D. Kim, C. Min, and S. Kang. Towards Automatic Recognition of Perceived Level of Understanding on Online Lectures using Earables. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 158–164, 2021.

[244] R. C. King, L. Atallah, C. Wong, F. Miskelly, and G. Yang. Elderly Risk Assessment of Falls with BSN. In *2010 International Conference*

on Body Sensor Networks, pages 30–35, June 2010. doi: 10.1109/BSN.2010.42.

[245] I. Klockhoff. Middle ear muscle reflexes in man. *Journal of Occupational and Environmental Medicine*, 3(10):486, 1961.

[246] M. T. Knierim, D. Puhl, G. Ivucic, and **T. Röddiger**. OpenBCI + 3D-Printed Headphones = Open ExG Headphones – An Open-Source Research Platform for Biopotential Earable Applications. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–7, 2023.

[247] J. F. Knight and C. Baber. A tool to assess the comfort of wearable computers. *Human factors*, 47(1):77–91, 2005.

[248] M. Koelle, S. Ananthanarayan, and S. Boll. Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–19, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376162. URL `https://doi.org/10.1145/3313831.3376162`.

[249] D. Kollmann, W. K. Hogan, C. Steidl, M. K. Hibbs-Brenner, D. S. Hedin, and P. A. Lichter. VCSEL based, wearable, continuously monitoring pulse oximeter. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4156–4159, July 2013. doi: 10.1109/EMBC.2013.6610460.

[250] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. H. Flowers, N. Miner, and J. Neuhoff. Sonification report: Status of the field and research agenda. 2010.

[251] E. Kristal-Boneh, P. Froom, G. Harari, M. Malik, and J. Ribak. Summer-winter differences in 24 h variability of heart rate. *European Journal of Cardiovascular Prevention & Rehabilitation*, 7(2):141–146, 2000.

[252] J. Kruger, H. M. Blanck, and C. Gillespie. Dietary and physical activity behaviors among adults successful at weight loss maintenance. *International Journal of Behavioral Nutrition and Physical Activity*, 3(1):1–10, 2006.

[253] K. Kunze, M. Katsutoshi, Y. Uema, and M. Inami. How much do you read? counting the number of words a user reads using electrooculography. In *Proceedings of the 6th Augmented Human International Conference*, pages 125–128, 2015.

[254] J. Laguarta, F. Hueto, and B. Subirana. COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings. *IEEE Open Journal of Engineering in Medicine and Biology*, 2020.

[255] D. Lambden. The Complete Guide to Hearable Technology in 2020, Oct 2020. URL `https://www.clearliving.com/hearing/technology/hearables/`.

[256] M. Laporte, P. Baglat, S. Gashi, M. Gjoreski, S. Santini, and M. Langheinrich. Detecting Verbal and Non-Verbal Gestures Using Earables. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 165–170, 2021.

[257] G. Laput, X. A. Chen, and C. Harrison. SweepSense: Ad Hoc Configuration Sensing Using Reflected Swept-Frequency Ultrasonics. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, IUI '16, page 332–335, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341370. doi: 10.1145/2856767.2856812. URL `https://doi.org/10.1145/2856767.2856812`.

[258] S. F. LeBoeuf, M. E. Aumer, W. E. Kraus, J. L. Johnson, and B. Duscha. Earbud-based sensor for the assessment of energy expenditure, heart rate, and VO2max. *Medicine and science in sports and exercise*, 46(5):1046, 2014.

[259] D. Lee, Y. Lee, Y. Shin, and I. Oakley. Designing Socially Acceptable Hand-to-Face Input. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, page 711–723, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450359481. doi: 10.1145/3242587.3242642. URL `https://doi.org/10.1145/3242587.3242642`.

[260] J. Lee, H.-S. Yeo, M. Dhuliawala, J. Akano, J. Shimizu, T. Starner, A. Quigley, W. Woo, and K. Kunze. Itchy Nose: Discreet Gesture Interaction Using EOG Sensors in Smart Eyewear. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ISWC '17, page 94–97, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450351881. doi: 10.1145/3123021.3123060. URL `https://doi.org/10.1145/3123021.3123060`.

[261] J. H. Lee, S. M. Lee, H. J. Byeon, J. S. Hong, K. S. Park, and S.-H. Lee. CNT/PDMS-based canal-typed ear electrodes for inconspicuous EEG recording. *Journal of Neural Engineering*, 11(4):046014, jun 2014. doi: 10.1088/1741-2560/11/4/046014. URL `https://doi.org/10.1088/1741-2560/11/4/046014`.

[262] K.-R. Lee, W.-D. Chang, S. Kim, and C.-H. Im. Real-time "eye-writing" recognition using electrooculogram. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(1):37–48, 2016.

[263] M. Lee, S. Je, W. Lee, D. Ashbrook, and A. Bianchi. ActivEarring: Spatiotemporal Haptic Cues on the Ears. *IEEE Transactions on Haptics*, 12(4):554–562, Oct 2019. ISSN 2329-4051. doi: 10.1109/TOH.2019.2925799.

[264] S. Lee, C. Min, A. Montanari, A. Mathur, Y. Chang, J. Song, and F. Kawsar. Automatic Smile and Frown Recognition with Kinetic Earables. In *Proceedings of the 10th Augmented Human International Conference 2019*, AH2019, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450365475. doi: 10.1145/3311823.3311869. URL `https://doi.org/10.1145/3311823.3311869`.

[265] S. Lee, C. Min, A. Montanari, A. Mathur, Y. Chang, J. Song, and F. Kawsar. Automatic Smile and Frown Recognition with Kinetic Earables. In *Proceedings of the 10th Augmented Human International Conference 2019*, pages 1–4, 2019.

[266] S. M. Lee, J. H. Lee, and S. Lee. Skin-like electronics based on CNT/PDMS composite for long term and unconscious sensing of biosignals. In *2015 IEEE 15th International Conference on Nanotechnology (IEEE-NANO)*, pages 1155–1158. IEEE, 2015.

[267] W. Lee, X. Yang, H. Jung, I. Bok, C. Kim, O. Kwon, and H. You. Anthropometric analysis of 3D ear scans of Koreans and Caucasians for ear product design. *Ergonomics*, 61(11):1480–1495, 2018.

[268] J. Leong, W. Lu, K. Luk, and E. Karlberg. Kinematics of the chest cage and spine during breathing in healthy individuals and in patients with adolescent idiopathic scoliosis. *Spine*, 24(13):1310, 1999.

[269] G. Li, Z. Zhang, and G. Wang. Emotion recognition based on low-cost in-ear EEG. In *2017 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, Oct 2017. doi: 10.1109/BIOCAS.2017.8325198.

[270] P. LIMITED. Thora-3Di Model T-01 510(k) Premarket Notification, Mar 2016. URL `https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K151940`.

[271] R. Lissermann, J. Huber, A. Hadjakos, and M. Mühlhäuser. Earput: Augmenting behind-the-ear devices for ear-based interaction. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 1323–1328. 2013.

[272] R. Lissermann, J. Huber, A. Hadjakos, S. Nanayakkara, and M. Mühlhäuser. EarPut: Augmenting Ear-Worn Devices for Ear-Based Interaction. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, OzCHI '14, page 300–307, New York, NY, USA, 2014. Association for

Computing Machinery. ISBN 9781450306539. doi: 10.1145/2686612.2686655. URL https://doi.org/10.1145/2686612.2686655.

[273] J. Liu, E. Johns, L. Atallah, C. Pettitt, B. Lo, G. Frost, and G. Yang. An Intelligent Food-Intake Monitoring System Using Wearable Sensors. In *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks*, pages 154–160, May 2012. doi: 10.1109/BSN.2012.11.

[274] J. Liu, W. Song, L. Shen, J. Han, X. Xu, and K. Ren. MandiPass: Secure and Usable User Authentication via Earphone IMU. In *2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)*, pages 674–684. IEEE, 2021.

[275] M. Liu, M. Nancel, and D. Vogel. Gunslinger: Subtle Arms-down Mid-Air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, page 63–71, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450337793. doi: 10.1145/2807442.2807489. URL https://doi.org/10.1145/2807442.2807489.

[276] Y. Liu and D. Hatzinakos. Earprint: Transient Evoked Otoacoustic Emission for Biometrics. *IEEE Transactions on Information Forensics and Security*, 9(12):2291–2301, Dec 2014. ISSN 1556-6021. doi: 10.1109/TIFS.2014.2361205.

[277] D. Looney, C. Park, P. Kidmose, M. L. Rank, M. Ungstrup, K. Rosenkranz, and D. P. Mandic. An in-the-ear platform for recording electroencephalogram. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6882–6885, Aug 2011. doi: 10.1109/IEMBS.2011.6091733.

[278] D. Looney, P. Kidmose, C. Park, M. Ungstrup, M. L. Rank, K. Rosenkranz, and D. P. Mandic. The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring. *IEEE Pulse*, 3(6):32–42, Nov 2012. ISSN 2154-2317. doi: 10.1109/MPUL.2012.2216717.

[279] D. Looney, P. Kidmose, and D. P. Mandic. Ear-EEG: user-centered and wearable BCI. In *Brain-Computer Interface Research*, pages 41–50. Springer, 2014.

[280] P. Lorenzi, R. Rao, G. Romano, A. Kita, and F. Irrera. Mobile Devices for the Real-Time Detection of Specific Human Motion Disorders. *IEEE Sensors Journal*, 16(23):8220–8227, Dec 2016. ISSN 1558-1748. doi: 10.1109/JSEN.2016.2530944.

[281] R. Lotfi, G. Tzanetakis, R. Eskicioglu, and P. Irani. A Comparison between Audio and IMU Data to Detect Chewing Events Based on an Earable Device. In *Proceedings of the 11th Augmented Human International Conference*, AH '20, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450377287. doi: 10.1145/3396339.3396362. URL `https://doi.org/10.1145/3396339.3396362`.

[282] R. Lotfi, G. Tzanetakis, R. Eskicioglu, and P. Irani. A comparison between audio and IMU data to detect chewing events based on an earable device. In *Proceedings of the 11th Augmented Human International Conference*, pages 1–8, 2020.

[283] M. Lueken, X. Feng, B. Venema, B. J. E. Misgeld, and S. Leonhardt. Photoplethysmography-based in-ear sensor system for identification of increased stress arousal in everyday life. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 83–86, May 2017. doi: 10.1109/BSN.2017.7936013.

[284] L. Luo, X. She, J. Cao, Y. Zhang, Y. Li, and P. X. K. Song. Detection and Prediction of Ovulation From Body Temperature Measured by an In-Ear Wearable Thermometer. *IEEE Transactions on Biomedical Engineering*, 67(2):512–522, Feb 2020. ISSN 1558-2531. doi: 10.1109/TBME.2019.2916823.

[285] D. Lupton. *The quantified self.* John Wiley & Sons, 2016.

[286] G. Ma, W. Zhu, J. Zhong, T. Tong, J. Zhang, and L. Wang. Wearable Ear Blood Oxygen Saturation and Pulse Measurement System Based on

PPG. In *2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 111–116, Oct 2018. doi: 10.1109/SmartWorld.2018.00054.

[287] B. Maag, Z. Zhou, O. Saukh, and L. Thiele. BARTON: Low Power Tongue Movement Sensing with In-Ear Barometers. In *2017 IEEE 23rd International Conference on Parallel and Distributed Systems (ICPADS)*, pages 9–16, Dec 2017. doi: 10.1109/ICPADS.2017.00013.

[288] S. Mahto, T. Arakawa, and T. Koshinak. Ear Acoustic Biometrics Using Inaudible Signals and Its Application to Continuous User Authentication. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1407–1411, Sep. 2018. doi: 10.23919/EUSIPCO.2018.8553015.

[289] H. Manabe and M. Fukumoto. Full-Time Wearable Headphone-Type Gaze Detector. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, page 1073–1078, New York, NY, USA, 2006. Association for Computing Machinery. ISBN 1595932984. doi: 10.1145/1125451.1125655. URL `https://doi.org/10.1145/1125451.1125655`.

[290] H. Manabe and M. Fukumoto. Tap Control for Headphones without Sensors. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, page 309–314, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450307161. doi: 10.1145/2047196.2047236. URL `https://doi.org/10.1145/2047196.2047236`.

[291] H. Manabe and M. Fukumoto. Tap control for headphones without sensors. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 309–314, 2011.

[292] H. Manabe, M. Fukumoto, and T. Yagi. Conductive Rubber Electrodes for Earphone-Based Eye Gesture Input Interface. In *Proceedings of*

the 2013 International Symposium on Wearable Computers, ISWC '13, page 33–40, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450321273. doi: 10.1145/2493988.2494329. URL `https://doi.org/10.1145/2493988.2494329`.

[293] S. Mandekar, L. Jentsch, D. K. Lutz, D. M. Behbahani, and M. Melnykowycz. Earable design analysis for sleep EEG measurements. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 171–175, 2021.

[294] N. Manivannan, N. Celik, and W. Balachandran. Evaluation of a behind-the-ear ECG device for smartphone based integrated multiple smart sensor system in health applications. 2016.

[295] A. Marinou, R. Saunders, and A. J. Casson. Flexible inkjet printed sensors for behind-the-ear SSVEP EEG monitoring. In *2020 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)*, pages 1–4, Aug 2020. doi: 10.1109/FLEPS49123.2020.9239488.

[296] A. Martin and J. Voix. In-Ear Audio Wearable: Measurement of Heart and Breathing Rates for Health and Safety Monitoring. *IEEE Transactions on Biomedical Engineering*, 65(6):1256–1263, June 2018. ISSN 1558-2531. doi: 10.1109/TBME.2017.2720463.

[297] M. Martinez and R. Stiefelhagen. Breath rate monitoring during sleep using near-IR imagery and PCA. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 3472–3475. IEEE, 2012.

[298] M. Masè, A. Micarelli, and G. Strapazzon. Hearables: new perspectives and pitfalls of in-ear devices for physiological monitoring. A scoping review. *Frontiers in Physiology*, 11:1227, 2020.

[299] M. J. Mathie, A. C. Coster, N. H. Lovell, and B. G. Celler. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological measurement*, 25(2):R1, 2004.

[300] K. Matsumoto, Y. Temiz, H. Taghavi, E. L. Cornelius, H. Mori, and B. Michel. An earbud-type wearable (A hearable) with vital parameter sensors for early detection and prevention of heat-stroke. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7049–7055, July 2019. doi: 10.1109/EMBC.2019.8856821.

[301] K. Matsumura and K. Okada. eSense Veers: A Case Study of Acoustical Manipulation in Walking without Sight both on Subtle and Overt Conditions. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 20–25, 2019.

[302] K. Matsumura, D. Sakamoto, M. Inami, and T. Igarashi. Universal Earphones: Earphones with Automatic Side and Shared Use Detection. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, page 305–306, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450310482. doi: 10.1145/2166966.2167025. URL `https://doi.org/10.1145/2166966.2167025`.

[303] S. Matsushita, T. Oba, K. Otsuki, M. Toji, J. Otsuki, and K. Ogawa. A wearable sense of balance monitoring system towards daily health care monitoring. In *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings.*, pages 176–183, Oct 2003. doi: 10.1109/ISWC.2003.1241409.

[304] D. J. C. Matthies, B. A. Strecker, and B. Urban. EarFieldSensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input through Facial Expressions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 1911–1922, New York, NY, USA, 2017. Association for Computing

Machinery. ISBN 9781450346559. doi: 10.1145/3025453.3025692. URL `https://doi.org/10.1145/3025453.3025692`.

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

[306] E. D. McCoul, V. K. Anand, and P. J. Christos. Validating the clinical assessment of eustachian tube dysfunction: The eustachian tube dysfunction questionnaire (ETDQ-7). *The Laryngoscope*, 122(5):1137–1141, 2012.

[307] D. McIlwraith, J. Pansiot, and G. Yang. Wearable and ambient sensor fusion for the characterisation of human motion. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5505–5510, Oct 2010. doi: 10.1109/IROS.2010.5650512.

[308] K. A. McMichael, J. Vander Bilt, L. Lavery, E. Rodriguez, and M. Ganguli. Simple balance and mobility tests can assess falls risk when cognition is impaired. *Geriatric Nursing*, 29(5):311–323, 2008.

[309] R. E. Mehler. *How the circulatory system works*. John Wiley & Sons, 2014.

[310] C. A. Merck, C. Maher, M. Mirtchouk, M. Zheng, Y. Huang, and S. Kleinberg. Multimodality sensing for eating recognition. In *PervasiveHealth*, pages 130–137, 2016.

[311] N. Merrill, M. T. Curran, J. Yang, and J. Chuang. Classifying mental gestures with in-ear EEG. In *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 130–135, June 2016. doi: 10.1109/BSN.2016.7516246.

[312] C. Metzger, M. Anderson, and T. Starner. FreeDigiter: a contact-free device for gesture control. In *Eighth International Symposium on Wearable Computers*, volume 1, pages 18–21, Oct 2004. doi: 10.1109/ISWC.2004.23.

[313] C. Metzger, M. Anderson, and T. Starner. Freedigiter: A contact-free device for gesture control. In *Eighth International Symposium on Wearable Computers*, volume 1, pages 18–21. IEEE, 2004.

[314] K. B. Mikkelsen, S. L. Kappel, D. P. Mandic, and P. Kidmose. EEG recorded from the ear: characterizing the ear-EEG method. *Frontiers in neuroscience*, 9:438, 2015.

[315] K. B. Mikkelsen, D. B. Villadsen, M. Otto, and P. Kidmose. Automatic sleep staging using ear-EEG. *Biomedical engineering online*, 16(1):1–15, 2017.

[316] K. B. Mikkelsen, S. L. Kappel, M. C. Hemmsen, M. L. Rank, and P. Kidmose. Discrimination of Sleep Spindles in Ear-EEG. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6697–6700. IEEE, 2019.

[317] K. B. Mikkelsen, Y. R. Tabar, S. L. Kappel, C. B. Christensen, H. O. Toft, M. C. Hemmsen, M. L. Rank, M. Otto, and P. Kidmose. Accurate whole-night sleep monitoring with dry-contact ear-EEG. *Scientific reports*, 9(1):1–12, 2019.

[318] C. Min, A. Mathur, and F. Kawsar. Exploring Audio and Kinetic Sensing on Earable Devices. In *Proceedings of the 4th ACM Workshop on Wearable Systems and Applications*, WearSys '18, page 5–10, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450358422. doi: 10.1145/3211960.3211970. URL `https://doi.org/10.1145/3211960.3211970`.

[319] C. Min, A. Mathur, and F. Kawsar. Audio-Kinetic Model for Automatic Dietary Monitoring with Earable Devices. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pages 517–517, 2018.

[320] C. Min, A. Mathur, and F. Kawsar. Exploring audio and kinetic sensing on earable devices. In *Proceedings of the 4th ACM Workshop on Wearable Systems and Applications*, pages 5–10, 2018.

[321] C. Min, A. Montanari, A. Mathur, S. Lee, and F. Kawsar. Cross-modal approach for conversational well-being monitoring with multi-sensory earables. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 706–709, 2018.

[322] C. Min, A. Mathur, A. Montanari, and F. Kawsar. An Early Characterisation of Wearing Variability on Motion Signals for Wearables. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC '19, page 166–168, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368704. doi: 10.1145/3341163.3347716. URL `https://doi.org/10.1145/3341163.3347716`.

[323] C. Min, A. Mathur, A. Montanari, and F. Kawsar. An early characterisation of wearing variability on motion signals for wearables. In *Proceedings of the 23rd International Symposium on Wearable Computers*, pages 166–168, 2019.

[324] M. Miravitlles and A. Ribera. Understanding the impact of symptoms on the burden of COPD. *Respiratory research*, 18(1):1–11, 2017.

[325] U. Mogera, A. A. Sagade, S. J. George, and G. U. Kulkarni. Ultrafast response humidity sensor using supramolecular nanofibre and its application in monitoring breath humidity and flow. *Scientific reports*, 4: 4103, 2014.

[326] A. R. Møller. Network model of the middle ear. *The Journal of the Acoustical Society of America*, 33(2):168–176, 1961.

[327] A. R. Møller. The acoustic middle ear muscle reflex. In *Auditory system*, pages 519–548. Springer, 1974.

[328] J. Monge-Álvarez, C. Hoyos-Barceló, P. Lesso, and P. Casaseca-de-la-Higuera. Robust Detection of Audio-Cough Events Using Local Hu Moments. *IEEE Journal of Biomedical and Health Informatics*, 23(1): 184–196, 2019.

[329] A. S. Monto, S. Gravenstein, M. Elliott, M. Colopy, and J. Schweinle. Clinical signs and symptoms predicting influenza infection. *Archives of internal medicine*, 160(21):3243–3247, 2000.

[330] N. Motokawa, A. Jinno, Y. Takayama, S. Ishii, A. Yokokubo, and G. Lopez. Coremoni-WE: Individual Core Training Monitoring and Support System Using an IMU at the Waist and the Ear. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 176–177, 2021.

[331] F. M. Munshi, E. S. A. El Zayat, and D. H. Dolmans. Development and utility of a questionnaire to evaluate the quality of PBL problems. *South East Asian Journal of Medical Education*, 2(2):32–40, 2008.

[332] R. C. Murdock and J. A. Hagen. Soldier safety and performance through wearable devices. In *Micro-and Nanotechnology Sensors, Systems, and Applications X*, volume 10639, page 106391A. International Society for Optics and Photonics, 2018.

[333] M. G. Myers and D. P. Olson. Central nervous system control of metabolism. *Nature*, 491(7424):357–363, 2012.

[334] R. Näätänen. Mismatch negativity: clinical research and possible applications. *International Journal of Psychophysiology*, 48(2):179–188, 2003.

[335] T. Nakamura, Y. D. Alqurashi, M. J. Morrell, and D. P. Mandic. Automatic detection of drowsiness using in-ear EEG. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, July 2018. doi: 10.1109/IJCNN.2018.8489723.

[336] T. Nakamura, V. Goverdovsky, and D. P. Mandic. In-Ear EEG Biometrics for Feasible and Readily Collectable Real-World Person Authentication. *IEEE Transactions on Information Forensics and Security*, 13 (3):648–661, March 2018. ISSN 1556-6021. doi: 10.1109/TIFS.2017.2763124.

[337] T. Nakamura, Y. D. Alqurashi, M. J. Morrell, and D. P. Mandic. Hearables: Automatic Overnight Sleep Monitoring With Standardized In-Ear EEG Sensor. *IEEE Transactions on Biomedical Engineering*, 67 (1):203–212, Jan 2020. ISSN 1558-2531. doi: 10.1109/TBME.2019.2911423.

[338] A. Nasser, K. Zhu, and S. Wiseman. Thermo-Haptic Earable Display for the Hearing and Visually Impaired. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '19, page 630–632, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366762. doi: 10.1145/3308561.3354636. URL `https://doi.org/10.1145/3308561.3354636`.

[339] A. Nasser, K. Zheng, and K. Zhu. ThermEarhook: Investigating Spatial Thermal Haptic Feedback on the Auricular Skin Area. In *Proceedings of the 2021 International Conference on Multimodal Interaction*, pages 662–672, 2021.

[340] C. K. H. Ne, J. Muzaffar, A. Amlani, and M. Bance. Hearables, in-ear sensing devices for bio-signal acquisition: a narrative review. *Expert Review of Medical Devices*, 18(sup1):95–128, 2021. doi: 10.1080/17434440.2021.2014321. URL `https://doi.org/10.1080/17434440.2021.2014321`. PMID: 34904507.

[341] E. Nemati, S. Zhang, T. Ahmed, M. M. Rahman, J. Kuang, and A. Gao. Coughbuddy: Multi-modal cough event detection using earbuds platform. In *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–4. IEEE, 2021.

[342] A. Nguyen, R. Alqurashi, Z. Raghebi, F. Banaei-kashani, A. C. Halbower, and T. Vu. A Lightweight and Inexpensive In-Ear Sensing System For Automatic Whole-Night Sleep Stage Monitoring. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, SenSys '16, page 230–244, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342636.

doi: 10.1145/2994551.2994562. URL `https://doi.org/10.1145/2994551.2994562`.

[343] N. Nguyen, A. Chakma, and N. Roy. A Scalable and Domain Adaptive Respiratory Symptoms Detection Framework using Earables. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 5620–5625. IEEE, 2021.

[344] P. Nguyen, N. Bui, A. Nguyen, H. Truong, A. Suresh, M. Whitlock, D. Pham, T. Dinh, and T. Vu. TYTH-Typing On Your Teeth: Tongue-Teeth Localization for Human-Computer Interface. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '18, page 269–282, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450357203. doi: 10.1145/3210240.3210322. URL `https://doi.org/10.1145/3210240.3210322`.

[345] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao. MusicalHeart: A <i>Hearty</i> Way of Listening to Music. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, SenSys '12, page 43–56, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450311694. doi: 10.1145/2426656.2426662. URL `https://doi.org/10.1145/2426656.2426662`.

[346] J. Nishimura and Tadahiro Kuroda. Eating habits monitoring using wireless wearable in-ear microphone. In *2008 3rd International Symposium on Wireless Pervasive Computing*, pages 130–132, May 2008. doi: 10.1109/ISWPC.2008.4556181.

[347] R. G. Norman, M. M. Ahmed, J. A. Walsleben, and D. M. Rapoport. Detection of respiratory events during NPSG: nasal cannula/pressure sensor versus thermistor. *Sleep*, 20(12):1175–1184, 1997.

[348] J. J. S. Norton, D. S. Lee, J. W. Lee, W. Lee, O. Kwon, P. Won, S.-Y. Jung, H. Cheng, J.-W. Jeong, A. Akce, S. Umunna, I. Na, Y. H.

Kwon, X.-Q. Wang, Z. Liu, U. Paik, Y. Huang, T. Bretl, W.-H. Yeo, and J. A. Rogers. Soft, curved electrode systems capable of integration on the auricle as a persistent brain-computer interface. *Proceedings of the National Academy of Sciences*, 112(13):3920–3925, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1424875112. URL `https://www.pnas.org/content/112/13/3920`.

[349] D. Nunan, G. R. Sandercock, and D. A. Brodie. A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. *Pacing and clinical electrophysiology*, 33(11):1407–1417, 2010.

[350] H. Odoemelem, A. Hölzemann, and K. Van Laerhoven. Using the ESense Wearable Earbud as a Light-Weight Robot Arm Controller. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 26–29, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369022. doi: 10.1145/3345615.3361138. URL `https://doi.org/10.1145/3345615.3361138`.

[351] H. Odoemelem, A. Hölzemann, and K. Van Laerhoven. Using the eSense Wearable Earbud as a Light-Weight Robot Arm Controller. 2019.

[352] C. J. O'Donnell and W. B. Kannel. Cardiovascular risks of hypertension: lessons from observational studies. *Journal of hypertension. Supplement: Official Journal of the International Society of Hypertension*, 16(6):S3–7, 1998.

[353] U. D. of Health, H. S. Food, D. A. C. for Drug Evaluation, R. C. Food, and D. Administration. Guidance for Industry, Investigating Out-of-Specification (OOS), Test Results for Pharmaceutical Production, 2006. URL `https://www.fda.gov/downloads/drugs/guidances/ucm070287.pdf`. [2018-12-02].

[354] R. J. Oliveira and G. Hoeker. Ear canal anatomy and activity. In *Seminars in hearing*, volume 24, pages 265–276. Copyright© 2003 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New . . . , 2003.

[355] K. Onizuka and C. G. Sodini. Head ballistocardiogram based on wireless multi-location sensors. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1275–1278, Aug 2015. doi: 10.1109/EMBC.2015.7318600.

[356] Z. Ouyang, J. Hu, J. Niu, and Z. Qi. An Asymmetrical Acoustic Field Detection System for Daily Tooth Brushing Monitoring. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6, Dec 2017. doi: 10.1109/GLOCOM.2017.8255017.

[357] M. A. Ozdemir, M. Degirmenci, E. Izci, and A. Akan. EEG-based emotion recognition with deep convolutional neural networks. *Biomedical Engineering/Biomedizinische Technik*, 66(1):43–57, 2021.

[358] J. Pansiot, R. C. King, D. G. McIlwraith, B. P. L. Lo, and Guang-Zhong Yang. ClimBSN: Climber performance monitoring with BSN. In *2008 5th International Summer School and Symposium on Medical Devices and Biosensors*, pages 33–36, June 2008. doi: 10.1109/ISSMDBS.2008.4575009.

[359] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos. The SPLENDID chewing detection challenge. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 817–820, July 2017. doi: 10.1109/EMBC.2017.8036949.

[360] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos. A Novel Chewing Detection System Based on PPG, Audio, and Accelerometry. *IEEE Journal of Biomedical and Health Informatics*, 21(3):607–618, May 2017. ISSN 2168-2208. doi: 10.1109/JBHI.2016.2625271.

[361] C. Park, C. Min, S. Bhattacharya, and F. Kawsar. Augmenting Conversational Agents with Ambient Acoustic Contexts. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–9, 2020.

[362] J.-H. Park, D.-G. Jang, J. W. Park, and S.-K. Youm. Wearable Sensing of In-Ear Pressure for Heart Rate Monitoring with a Piezoelectric Sensor. *Sensors*, 15(9):23402–23417, 2015. ISSN 1424-8220. doi: 10.3390/s150923402. URL `https://www.mdpi.com/1424-8220/15/9/23402`.

[363] R. Parkes. Rate of respiration: the forgotten vital sign. *Emergency Nurse*, 19(2), 2011.

[364] S. Paßler and W. Fischer. Food Intake Activity Detection Using a Wearable Microphone System. In *2011 Seventh International Conference on Intelligent Environments*, pages 298–301, July 2011. doi: 10.1109/IE.2011.9.

[365] S. Paßler and W. Fischer. Food Intake Activity Detection Using a Wearable Microphone System. In *2011 Seventh International Conference on Intelligent Environments*, pages 298–301, July 2011. doi: 10.1109/IE.2011.9.

[366] S. Päßler, M. Wolff, and W.-J. Fischer. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. *Physiological Measurement*, 33(6): 1073–1093, may 2012. doi: 10.1088/0967-3334/33/6/1073. URL `https://doi.org/10.1088/0967-3334/33/6/1073`.

[367] S. Passler, N. Mü"ller, and V. Senner. In-Ear Pulse Rate Measurement: A Valid Alternative to Heart Rate Derived from Electrocardiography? *Sensors*, 19(17), 2019. ISSN 1424-8220. doi: 10.3390/s19173641. URL `https://www.mdpi.com/1424-8220/19/17/3641`.

[368] S. Passler, N. Müller, and V. Senner. In-Ear Pulse Rate Measurement: A Valid Alternative to Heart Rate Derived from Electrocardiography? *Sensors*, 19(17):3641, 2019.

[369] A. K. Patel, V. Reddy, and J. F. Araujo. Physiology, sleep stages. *StatPearls [Internet]*, 2021.

[370] A. Paul, S. R. Deiss, D. Tourtelotte, M. Kleffner, T. Zhang, and G. Cauwenberghs. Electrode-Skin Impedance Characterization of In-Ear Electrophysiology Accounting for Cerumen and Electrodermal Response. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 855–858, March 2019. doi: 10.1109/NER.2019.8716918.

[371] A. Paul, G. Hota, B. Khaleghi, Y. Xu, T. Rosing, and G. Cauwenberghs. Attention State Classification with In-Ear EEG. In *2021 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–5. IEEE, 2021.

[372] A. Pedrana, D. Comotti, V. Re, and G. Traversi. Development of a Wearable In-Ear PPG System for Continuous Monitoring. *IEEE Sensors Journal*, 20(23):14482–14490, Dec 2020. ISSN 1558-1748. doi: 10.1109/JSEN.2020.3008479.

[373] F. Pfreundtner, J. Yang, and G. Sörös. (W) Earable Microphone Array and Ultrasonic Echo Localization for Coarse Indoor Environment Mapping. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4475–4479. IEEE, 2021.

[374] N. Pham, T. Dinh, Z. Raghebi, T. Kim, N. Bui, P. Nguyen, H. Truong, F. Banaei-Kashani, A. Halbower, T. Dinh, and T. Vu. WAKE: A behind-the-Ear Wearable System for Microsleep Detection. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, MobiSys '20, page 404–418, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379540. doi: 10.1145/3386901.3389032. URL `https://doi.org/10.1145/3386901.3389032`.

[375] N. Pham, T. Dinh, Z. Raghebi, T. Kim, N. Bui, P. Nguyen, H. Truong, F. Banaei-Kashani, A. Halbower, T. Dinh, et al. WAKE: a behind-the-ear wearable system for microsleep detection. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pages 404–418, 2020.

[376] D. Phan, L. Y. Siong, P. N. Pathirana, and A. Seneviratne. Smartwatch: Performance evaluation for long-term heart rate monitoring. In *2015 International symposium on bioelectronics and bioinformatics (ISBB)*, pages 144–147. IEEE, 2015.

[377] R. W. Picard, S. Fedor, and Y. Ayzenberg. Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, 8(1):62–75, 2016.

[378] C. Pirzanski and B. Berge. Ear canal dynamics: Facts versus perception. *The Hearing Journal*, 58(10):50–52, 2005.

[379] J. Plazak and M. Kersten-Oertel. A Survey on the Affordances of "Hearables". *Inventions*, 3(3):48, 2018.

[380] R. Plutchik. Emotions in the practice of psychotherapy-clinical implications of affect theories. 2000.

[381] M. Poh, K. Kim, A. D. Goessling, N. C. Swenson, and R. W. Picard. Heartphones: Sensor Earphones and Mobile Application for Non-obtrusive Health Monitoring. In *2009 International Symposium on Wearable Computers*, pages 153–154, Sep. 2009. doi: 10.1109/ISWC.2009.35.

[382] M. Poh, N. C. Swenson, and R. W. Picard. Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. *IEEE Transactions on Information Technology in Biomedicine*, 14 (3):786–794, May 2010. ISSN 1558-0032. doi: 10.1109/TITB.2010.2042607.

[383] H. Pohl, A. Muresan, and K. Hornbæk. Charting Subtle Interaction in the HCI Literature. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–15, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. doi: 10.1145/3290605.3300648. URL `https://doi.org/10.1145/3290605.3300648`.

[384] A. Politzer. *A Text-Book of the Diseases of the Ear for Students and Practitioners*. Baillière, Tindall and Cox, 1902.

[385] A. Politzer. *A Text-book of the Diseases of the Ear: For Students and Practitioners*. Lea Bros., 1909.

[386] J. Powar and A. R. Beresford. A Data Sharing Platform for Earables Research. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 30–35, 2019.

[387] J. Prakash, Z. Yang, Y.-L. Wei, and R. R. Choudhury. STEAR: Robust Step Counting from Earables. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 36–41, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369022. doi: 10.1145/3345615.3361133. URL `https://doi.org/10.1145/3345615.3361133`.

[388] J. Prakash, Z. Yang, Y.-L. Wei, and R. R. Choudhury. STEAR: Robust Step Counting from Earables. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 36–41, 2019.

[389] J. Prakash, Z. Yang, Y.-L. Wei, H. Hassanieh, and R. R. Choudhury. EarSense: Earphones as a Teeth Activity Sensor. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, MobiCom '20, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450370851. doi: 10.1145/3372224.3419197. URL `https://doi.org/10.1145/3372224.3419197`.

[390] G. A. Pressler, J. P. Mansfield, H. Pasterkamp, and G. R. Wodicka. Detection of respiratory sounds at the external ear. *IEEE transactions on biomedical engineering*, 51(12):2089–2096, 2004.

[391] S. A. Purabi, R. Rashed, M. Islam, N. Uddin, M. Naznin, and A. A. A. Islam. As you are, so shall you move your head: a system-level analysis between head movements and corresponding traits and emotions. In *Proceedings of the 6th International Conference on Networking, Systems and Security*, pages 3–11, 2019.

[392] H.-K. Ra, J. Ahn, H. J. Yoon, D. Yoon, S. H. Son, and J. Ko. I am a" smart" watch, smart enough to know the accuracy of my own heart rate sensor. In *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, pages 49–54, 2017.

[393] M. Radhakrishnan and A. Misra. Can Earables Support Effective User Engagement during Weight-Based Gym Exercises? In *Proceedings of the 1st International Workshop on Earable Computing*, pages 42–47, 2019.

[394] M. Radhakrishnan, D. Rathnayake, O. K. Han, I. Hwang, and A. Misra. ERICA: enabling real-time mistake detection & corrective feedback for free-weights exercises. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 558–571, 2020.

[395] M. Radhakrishnan, K. Misra, and V. Ravichandran. Applying "Earable" Inertial Sensing for Real-time Head Posture Detection. In *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 176–181. IEEE, 2021.

[396] M. D. Reilly, H. Shen, P. R. Calder, and H. B.-L. Duh. Understanding the Effects of Discreet Real-Time Social Interaction on Student Engagement in Lectures. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, OzCHI '13, page 193–196, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450325257. doi: 10.1145/2541016.2541086. URL `https://doi.org/10.1145/2541016.2541086`.

[397] G. R. Research, Jul 2020. URL `https://www.grandviewresearch.com/industry-analysis/earphone-and-headphone-market`.

[398] J. Rico and S. Brewster. Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–2, 2009.

[399] P. Ritz and G. Berrut. The importance of good hydration for day-to-day health. *Nutrition reviews*, 63(suppl_1):S6–S13, 2005.

[400] D. A. Robinson. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, 180(3):569, 1965.

[401] C. Rojas, N. Poulsen, M. Van Tuyl, D. Vargas, Z. Cohen, J. Paradiso, P. Maes, K. Esvelt, and F. Adib. A Scalable Solution for Signaling Face Touches to Reduce the Spread of Surface-based Pathogens. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–22, 2021.

[402] E. M. Roos and L. S. Lohmander. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health and quality of life outcomes*, 1(1):1–8, 2003.

[403] C. Roossien, R. Heus, M. Reneman, and G. Verkerke. Monitoring core temperature of firefighters to validate a wearable non-invasive core thermometer in different types of protective clothing: concurrent in-vivo validation. *Applied ergonomics*, 83:103001, 2020.

[404] S. Rupavatharam and M. Gruteser. Towards In-Ear Inertial Jaw Clenching Detection. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 54–55, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369022. doi: 10.1145/3345615.3361134. URL `https://doi.org/10.1145/3345615.3361134`.

[405] S. Rupavatharam and M. Gruteser. Towards in-ear inertial jaw clenching detection. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 54–55, 2019.

[406] J. A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[407] H. Sahni, A. Bedri, G. Reyes, P. Thukral, Z. Guo, T. Starner, and M. Ghovanloo. The Tongue and Ear Interface: A Wearable System for Silent Speech Recognition. In *Proceedings of the 2014 ACM*

International Symposium on Wearable Computers, ISWC '14, page 47–54, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450329699. doi: 10.1145/2634317.2634322. URL `https://doi.org/10.1145/2634317.2634322`.

[408] G. Salomon and A. Starr. Electromyography of middle ear muscles in man during motor activities. *Acta Neurologica Scandinavica*, 39(2): 161–168, 1963.

[409] R. S. Salzar, R. Cameron, and J. A. Pellettiere. Improving earpiece accelerometer coupling to the head. *SAE International Journal of Passenger Cars-Mechanical Systems*, 1(2008-01-2978):1367–1381, 2008.

[410] T. Sand, M. B. Kvaløy, T. Wader, and H. Hovdal. Evoked potential tests in clinical diagnosis. *Tidsskrift for Den norske legeforening*, 2013.

[411] A. Sano, T. Tomita, and H. Oba. Applications using earphone with biosignal sensors. In *Human Interface Society Meeting*, volume 12, pages 1–6. Citeseer, 2010.

[412] M. A. Sasse and C. Johnson. The perceptual window: Head motion as a new input stream. In *Human-computer Interaction, INTERACT'99: IFIP TC. 13 International Conference on Human-Computer Interaction, 30th August-3rd September 1999, Edinburgh, UK*, volume 1, page 238. IOS Press, 1999.

[413] J. Sauro. *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011.

[414] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E. L. Melanson, and M. Neuman. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiological Measurement*, 29(5):525–541, apr 2008. doi: 10.1088/0967-3334/29/5/001. URL `https://doi.org/10.1088/0967-3334/29/5/001`.

[415] E. S. Sazonov and J. M. Fontana. A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing. *IEEE Sensors Journal*, 12(5):1340–1348, May 2012. ISSN 1558-1748. doi: 10.1109/JSEN.2011.2172411.

[416] G. Schindler, C. Metzger, and T. Starner. A wearable interface for topological mapping and localization in indoor environments. In *International Symposium on Location-and Context-Awareness*, pages 64–73. Springer, 2006.

[417] N. Selvaraj. Assessment of pulse transit/arrival time as noninvasive blood pressure predictors in finger and earlobe sites. In *2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)*, pages 200–203, Nov 2016. doi: 10.1109/HIC.2016.7797731.

[418] A. Shamekhi and T. Bickmore. Breathe Deep: A Breath-Sensitive Interactive Meditation Coach. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 108–117. ACM, 2018.

[419] T. Shen, Tim Hsiao, Yu-Tsung Liu, and Tsung-Ying He. An ear-lead ECG based smart sensor system with voice biofeedback for daily activity monitoring. In *TENCON 2008 - 2008 IEEE Region 10 Conference*, pages 1–6, Nov 2008. doi: 10.1109/TENCON.2008.4766645.

[420] S. Shetty, V. Pitti, C. Satish Babu, G. Surendra Kumar, and B. Deepthi. Bruxism: a literature review. *The Journal of Indian prosthodontic society*, 10(3):141–148, 2010.

[421] M. Shibasaki, T. E. Wilson, and C. G. Crandall. Neural control and mechanisms of eccrine sweating during heat stress and exercise. *Journal of applied physiology*, 100(5):1692–1701, 2006.

[422] K. Shin, Y. Kim, S. Bae, K. Park, and S. Kim. A novel headset with a transmissive PPG sensor for heart rate measurement. In *13th International Conference on Biomedical Engineering*, pages 519–522. Springer, 2009.

[423] K. Shirota, R. L. Peiris, and K. Minamizawa. Altered Pinna: Exploring Shape Change of Pinna for Perception and Illusion of Sound Direction Change. In *Proceedings of the 23rd International Symposium on Wearable Computers*, ISWC '19, page 220–224, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368704. doi: 10.1145/3341163.3347725. URL `https://doi.org/10.1145/3341163.3347725`.

[424] M. Shuzo, S. Komori, T. Takashima, G. Lopez, S. Tatsuta, S. Yanagimoto, S. Warisawa, J.-J. Delaunay, and I. Yamada. Wearable eating habit sensing system using internal body sound. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 4(1):158–166, 2010.

[425] R. C. Simpson and H. H. Koester. Adaptive one-switch row-column scanning. *IEEE Transactions on Rehabilitation Engineering*, 7(4):464–473, 1999.

[426] C. J. Smith and G. Havenith. Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia. *European journal of applied physiology*, 111(7):1391–1404, 2011.

[427] J. Smith. Monitoring chronic cough: current and future techniques. *Expert review of respiratory medicine*, 4(5):673–683, 2010.

[428] H. Soronen, M. Turunen, and J. Hakulinen. Voice commands in home environment-a consumer survey. In *Ninth Annual Conference of the International Speech Communication Association*, 2008.

[429] Soundcore. Anker Soundcore Liberty 4. URL `https://de.soundcore.com/products/liberty-4-a3953g11`.

[430] C. Stampi, P. Stone, and A. Michimori. A new quantitative method for assessing sleepiness: the Alpha Attenuation Test. *Work & Stress*, 9(2-3): 368–376, 1995.

[431] S. Standring. *Gray's anatomy e-book: the anatomical basis of clinical practice*. Elsevier Health Sciences, 2015.

[432] A. Stochholm, K. Mikkelsen, and P. Kidmose. Automatic sleep stage classification using ear-EEG. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4751–4754, Aug 2016. doi: 10.1109/EMBC.2016.7591789.

[433] D. Strömbäck, S. Huang, and V. Radu. MM-Fit: Multimodal Deep Learning for Automatic Exercise Logging across Sensing Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–22, 2020.

[434] C. Sugimoto and R. Kohno. Development of wireless sensing system monitoring physiological information for healthcare in daily life. In *2011 Fifth International Conference on Sensing Technology*, pages 488–493, Nov 2011. doi: 10.1109/ICSensT.2011.6137028.

[435] W. Sun, F. M. Li, B. Steeper, S. Xu, F. Tian, and C. Zhang. TeethTap: Recognizing Discrete Teeth Gestures Using Motion and Acoustic Sensing on an Earpiece. In *26th International Conference on Intelligent User Interfaces*, pages 161–169, 2021.

[436] X. Sun, L. Qiu, Y. Wu, Y. Tang, and G. Cao. Sleepmonitor: Monitoring respiratory rate and body position during sleep using smartwatch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):104, 2017.

[437] S. Sur and V. K. Sinha. Event-related potential: An overview. *Industrial psychiatry journal*, 18(1):70, 2009.

[438] B. Surawicz and T. Knilans. *Chou's Electrocardiography in Clinical Practice E-Book: Adult and Pediatric*. Elsevier Health Sciences, 2008.

[439] S. Suzuki, Y. Tomita, and Y. Mitsukura. Stress and drowsiness evaluation BBased on earbud-type photoplethysmographic sensor. In *2015 10th Asian Control Conference (ASCC)*, pages 1–6, May 2015. doi: 10.1109/ASCC.2015.7244504.

[440] T. Sztyler and H. Stuckenschmidt. On-body localization of wearable devices: An investigation of position-aware activity recognition. In *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–9. IEEE, 2016.

[441] J. Taelman, S. Vandeput, A. Spaepen, and S. V. Huffel. Influence of mental stress on heart rate and heart rate variability. In *4th European conference of the international federation for medical and biological engineering*, pages 1366–1369. Springer, 2009.

[442] Y. Takayama, S. Ishii, A. Yokokubo, and G. Lopez. Detecting forward leaning posture using eSense and developing a posture improvement promoting system. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 178–179, 2021.

[443] E. Tamaki, T. Miyaki, and J. Rekimoto. Brainy Hand: An Ear-Worn Hand Gesture Interaction Device. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, page 4255–4260, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605582474. doi: 10.1145/1520340.1520649. URL `https://doi.org/10.1145/1520340.1520649`.

[444] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa. Earable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors*, 18(3), 2018. ISSN 1424-8220. doi: 10.3390/s18030733. URL `https://www.mdpi.com/1424-8220/18/3/733`.

[445] X. Teng and Y. Zhang. Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439)*, volume 4, pages 3153–3156. IEEE, 2003.

[446] M. Teplan et al. Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11, 2002.

[447] B. Tessendorf, P. Derleth, M. Feilner, F. Gravenhorst, A. Kettner, D. Roggen, T. Stiefmeier, and G. Tröster. Ear-worn reference data collection and annotation for multimodal context-aware hearing instruments. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2468–2471, Aug 2012. doi: 10.1109/EMBC.2012.6346464.

[448] **Röddiger, Tobias**, D. Wolffram, D. Laubenstein, M. Budde, and M. Beigl. Towards respiration rate monitoring using an in-ear headphone inertial measurement unit. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 48–53, 2019.

[449] **T. Röddiger**, M. Beigl, M. Köpke, and M. Budde. VOCNEA: sleep apnea and hypopnea detection using a novel tiny gas sensor. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, pages 226–227, 2018.

[450] **T. Röddiger**, D. Wolffram, D. Laubenstein, M. Budde, and M. Beigl. Towards Respiration Rate Monitoring Using an In-Ear Headphone Inertial Measurement Unit. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 48–53. Association for Computing Machinery, 2019.

[451] **T. Röddiger**, M. Beigl, and A. Exler. Design space and usability of earable prototyping. In *Proceedings of the 2020 International Symposium on Wearable Computers*, pages 73–78, 2020.

[452] **T. Röddiger**, M. Beigl, M. Hefenbrock, D. Wolffram, and E. Pescara. Detecting Episodes of Increased Cough Using Kinetic Earables. In *Augmented Humans Conference 2021*, pages 111–115, 2021.

[453] **T. Röddiger**, C. Clarke, D. Wolffram, M. Budde, and M. Beigl. EarRumble: Discreet Hands-and Eyes-Free Input by Voluntary Tensor

Tympani Muscle Contraction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[454] **T. Röddiger**, C. Dinse, and M. Beigl. Wearability and Comfort of Earables During Sleep. In *2021 International Symposium on Wearable Computers*, pages 150–152, 2021.

[455] **T. Röddiger**, C. Clarke, P. Breitling, T. Schneegans, H. Zhao, H. Gellersen, and M. Beigl. Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3):1–57, 2022.

[456] **T. Röddiger**, T. King, D. R. Roodt, C. Clarke, and M. Beigl. Open-Earable: Open Hardware Earable Sensing Platform. In *Proceedings of the 1st International Workshop on Earable Computing, EarComp*, volume 22, pages 29–34, 2022.

[457] J. F. Thayer and R. D. Lane. The role of vagal function in the risk for cardiovascular disease and mortality. *Biological psychology*, 74(2): 224–242, 2007.

[458] J. F. Thayer, S. S. Yamamoto, and J. F. Brosschot. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *International journal of cardiology*, 141(2):122–131, 2010.

[459] T. Tigges, T. Büchler, A. Pielmuş, M. Klum, A. Feldheiser, O. Hunsicker, and R. Orglmeister. Assessment of in-ear photoplethysmography as a surrogate for electrocardiography in heart rate variability analysis. In *World Congress on Medical Physics and Biomedical Engineering 2018*, pages 293–297. Springer, 2019.

[460] J. Tokuda and Kajiro Watanabe. Development of an earphone-like pulsimeter and walking application. In *2008 SICE Annual Conference*, pages 2304–2307, Aug 2008. doi: 10.1109/SICE.2008.4655049.

[461] Y. Tomita, S. Suzuki, and Y. Mitsukura. An earbud-based photoplethysmography and its application. *IEEJ Transactions on Electronics, Information and Systems*, 137(2):236–241, 2017.

[462] J. L. Tracy and D. Randles. Four models of basic emotions: a review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion review*, 3(4):397–405, 2011.

[463] O. Valentin, G. Viallet, A. Delnavaz, G. Cretot-Richert, M. Ducharme, H. Monsarat-Chanon, and J. Voix. Custom-fitted in-and around-the-ear sensors for unobtrusive and on-the-go EEG acquisitions: Development and validation. *Sensors*, 21(9):2953, 2021.

[464] M. Valentini and G. Parati. Variables influencing heart rate. *Progress in cardiovascular diseases*, 52(1):11–19, 2009.

[465] S. L. Vaught. Gait, balance, and fall prevention. *Ochsner Journal*, 3(2): 94–97, 2001.

[466] T. Vega Gálvez, S. Sapkota, A. Dancu, and P. Maes. Byte.It: Discreet Teeth Gestures for Mobile Device Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, page 1–6, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359719. doi: 10.1145/3290607.3312925. URL `https://doi.org/10.1145/3290607.3312925`.

[467] E. Velloso, M. Carter, J. Newn, A. Esteves, C. Clarke, and H. Gellersen. Motion correlation: Selecting objects by matching their movement. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(3): 1–35, 2017.

[468] B. Venema and S. Leonhardt. In-ear pulse oximetry in high altitude mountaineering. In *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, pages 254–259, Feb 2017. doi: 10.2316/P.2017.852-044.

[469] B. Venema, J. Schiefer, V. Blazek, N. Blanik, and S. Leonhardt. Evaluating Innovative In-Ear Pulse Oximetry for Unobtrusive Cardiovascular and Pulmonary Monitoring During Sleep. *IEEE Journal of Translational Engineering in Health and Medicine*, 1:2700208–2700208, 2013. ISSN 2168-2372. doi: 10.1109/JTEHM.2013.2277870.

[470] B. Venema, V. Blazek, and S. Leonhardt. In-ear photoplethysmography for mobile cardiorespiratory monitoring and alarming. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–5, June 2015. doi: 10.1109/BSN.2015.7299367.

[471] S. Venkatesh, C. R. Anderson, N. V. Rivera, and R. M. Buehrer. Implementation and analysis of respiration-rate estimation using impulse-based UWB. In *MILCOM 2005-2005 IEEE Military Communications Conference*, pages 3314–3320. IEEE, 2005.

[472] D. Verma, S. Bhalla, D. Sahnan, J. Shukla, and A. Parnami. ExpressEar: Sensing Fine-Grained Facial Expressions with Earables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(3):1–28, 2021.

[473] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 439–448, 2013.

[474] S. Vogel, M. Hulsbusch, D. Starke, and S. Leonhardt. In-Ear Heart Rate Monitoring Using a Micro-Optic Reflective Sensor. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1375–1378, Aug 2007. doi: 10.1109/IEMBS.2007.4352554.

[475] S. Vogel, M. Hulsbusch, D. Starke, and S. Leonhardt. A system for assessing motion artifacts in the signal of a micro-optic in-ear vital signs

sensor. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 510–513, Aug 2008. doi: 10.1109/IEMBS.2008.4649202.

[476] S. Vogel, M. Hü"lsbusch, T. Hennig, V. Blazek, and S. Leonhardt. In-Ear Vital Signs Monitoring Using a Novel Microoptic Reflective Sensor. *IEEE Transactions on Information Technology in Biomedicine*, 13 (6):882–889, Nov 2009. ISSN 1558-0032. doi: 10.1109/TITB.2009.2033268.

[477] W. von Rosenberg, T. Chanwimalueang, V. Goverdovsky, N. S. Peters, C. Papavassiliou, and D. P. Mandic. Hearables: Feasibility of recording cardiac rhythms from head and in-ear locations. *Royal Society open science*, 4(11):171214, 2017.

[478] C.-C. Wang, C.-Y. Chiang, P.-Y. Lin, Y.-C. Chou, I.-T. Kuo, C.-N. Huang, and C.-T. Chan. Development of a fall detecting system for the elderly residents. In *2008 2nd international conference on bioinformatics and biomedical engineering*, pages 1359–1362. IEEE, 2008.

[479] H. Wang, J. M. Lane, S. E. Jones, H. S. Dashti, H. M. Ollila, A. R. Wood, V. T. Van Hees, B. Brumpton, B. S. Winsvold, K. Kantojärvi, et al. Genome-wide association analysis of self-reported daytime sleepiness identifies 42 loci that suggest biological subtypes. *Nature communications*, 10(1):1–12, 2019.

[480] K. Wang, K. You, F. Chen, Z. Huang, and Z. Mao. Human-machine interface using eye saccade and facial expression physiological signals to improve the maneuverability of wearable robots. In *2017 International Symposium on Wearable Robotics and Rehabilitation (WeRob)*, pages 1–2, Nov 2017. doi: 10.1109/WEROB.2017.8383845.

[481] L. Wang, B. P. Lo, and G. Z. Yang. Multichannel Reflective PPG Earpiece Sensor With Passive Motion Cancellation. *IEEE Transactions on Biomedical Circuits and Systems*, 1(4):235–241, Dec 2007. ISSN 1940-9990. doi: 10.1109/TBCAS.2007.910900.

[482] Y. Wang, M. Nakanishi, S. L. Kappel, P. Kidmose, D. P. Mandic, Y. Wang, C. Cheng, and T. Jung. Developing an online steady-state visual evoked potential-based brain-computer interface system using EarEEG. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2271–2274, Aug 2015. doi: 10.1109/EMBC.2015.7318845.

[483] Z. Wang, S. Tan, L. Zhang, Y. Ren, Z. Wang, and J. Yang. EarDynamic: An Ear Canal Deformation Based Continuous User Authentication Using In-Ear Wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–27, 2021.

[484] Z. Wang, S. Tan, L. Zhang, Y. Ren, Z. Wang, and J. Yang. EarDynamic: An Ear Canal Deformation Based Continuous User Authentication Using In-Ear Wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–27, 2021.

[485] J. G. Webster. *Design of pulse oximeters*. CRC Press, 1997.

[486] M. Weigel, T. Lu, G. Bailly, A. Oulasvirta, C. Majidi, and J. Steimle. ISkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 2991–3000, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456. doi: 10.1145/2702123.2702391. URL `https://doi.org/10.1145/2702123.2702391`.

[487] J. Weisenberger, A. Heidbreder, and J. Miller. Development and preliminary evaluation of an earmold sound-to-tactile aid for the hearing-impaired. *J Rehabil Res Dev*, 24:51–66, 1987.

[488] D. Wendt, L. J. Van Loon, and W. D. M. Lichtenbelt. Thermoregulation during exercise in the heat. *Sports medicine*, 37(8):669–682, 2007.

[489] B. Wijam and Y. Wongsawat. An Investigation on Alpha-Beta Ratio of EarEEG System. In *2021 18th International Conference on Electrical*

Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pages 1117–1120. IEEE, 2021.

[490] E. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949.

[491] J. D. Witt, J. R. Fisher, J. A. Guenette, K. A. Cheong, B. J. Wilson, and A. W. Sheel. Measurement of exercise ventilation by a portable respiratory inductive plethysmograph. *Respiratory physiology & neurobiology*, 154(3):389–395, 2006.

[492] K. Wolf, A. Naumann, M. Rohs, and J. Müller. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-Dependent requirements. In *IFIP conference on human-computer interaction*, pages 559–575. Springer, 2011.

[493] O. J. Woodman. An introduction to inertial navigation. Technical report, University of Cambridge, Computer Laboratory, 2007.

[494] X. Xu, H. Shi, X. Yi, W. Liu, Y. Yan, Y. Shi, A. Mariakakis, J. Mankoff, and A. K. Dey. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–14, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376836. URL `https://doi.org/10.1145/3313831.3376836`.

[495] X. Xu, H. Shi, X. Yi, W. Liu, Y. Yan, Y. Shi, A. Mariakakis, J. Mankoff, and A. K. Dey. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[496] Y. Yan, C. Yu, X. Yi, and Y. Shi. Headgesture: hands-free input approach leveraging head movements for hmd devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2 (4):1–23, 2018.

[497] Y. Yan, C. Yu, Y. Shi, and M. Xie. PrivateTalk: Activating Voice Input with Hand-On-Mouth Gesture Detected by Bluetooth Earphones. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, page 1013–1020, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368162. doi: 10.1145/3332165.3347950. URL `https://doi.org/10.1145/3332165.3347950`.

[498] Yunjoo Lee, Hyonyoung Han, and Jung Kim. Influence of motion artifacts on photoplethysmographic signals for measuring pulse rates. In *2008 International Conference on Control, Automation and Systems*, pages 962–965, Oct 2008. doi: 10.1109/ICCAS.2008.4694635.

[499] C. Zeagler. Where to wear it: functional, technical, and social considerations in on-body location for wearable technology 20 years of designing for wearability. pages 150–157, 09 2017. doi: 10.1145/3123021.3123042.

[500] J. Zeilfelder, M. Diehl, C. Pylatiuk, and W. Stork. Concept for a Permanent, Non-Invasive Blood Pressure Measurement in the Ear. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 1–4. IEEE, 2019.

[501] G. Q. Zhang and W. Zhang. Heart rate, lifespan, and mortality risk. *Ageing research reviews*, 8(1):52–60, 2009. ISSN 1568-1637. doi: https://doi.org/10.1016/j.arr.2008.10.001. URL `https://www.sciencedirect.com/science/article/pii/S1568163708000524`.

[502] Q. Zhang, D. Zhou, and X. Zeng. Hear the heart: Daily cardiac health monitoring using Ear-ECG and machine learning. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, pages 448–451, Oct 2017. doi: 10.1109/UEMCON.2017.8249110.

[503] S. Zhang, E. Nemati, T. Ahmed, M. M. Rahman, J. Kuang, and A. Gao. A novel multi-centroid template matching algorithm and its application

to cough detection. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 7598–7604. IEEE, 2021.

[504] T. Zhang, X. Yi, R. Wang, Y. Wang, C. Yu, Y. Lu, and Y. Shi. Tap-to-Pair: Associating Wireless Devices with Synchronous Tapping. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4):1–21, 2018.

[505] H. Zhao, **T. Röddiger**, and M. Beigl. AirCase: Earable Charging Case with Air Quality Monitoring and Soundscape Sonification. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pages 180–184, 2021.

[506] X. Zhu, W.-L. Zheng, B.-L. Lu, X. Chen, S. Chen, and C. Wang. EOG-based drowsiness detection using Convolutional Neural Networks. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 128–134. IEEE, 2014.

A. Appendix

A.1 Heart Rate

Heart rate sensing with earables is summarized in subsubsection 3.1.1.1. The appendix Table A.1 lists the papers that developed heart rate sensing earables according to the sensor used, the activity level and context they have been evaluated with (e.g., while sleeping, speaking, etc.), where the sensor is located, and what was the achieved performance. Additionally, the reference device and number of study participants are listed.

Commonly, earable prototypes for heart sensing are evaluated in comparison to a gold-standard ECG or PPG on other parts of the body (e.g., a fingertip PPG [345; 209] or chest-ECG [422; 173]). Most studies cannot be directly compared because of different sensors, experimental contexts, and performance metrics. Performance is assessed based on correlations (e.g., [417; 476]) mean or median deviations (e.g., [345; 476; 474]), or standard deviations (e.g., [469; 470]). Several studies applied more unconventional metrics. For example, Wang et al. [481] measure the accuracy of the movement detection and the effects of motion artifacts with the HR spectrum fidelity index (fHRS), which is the ratio of the HR frequency spectrum density over the full-spectrum strength for their respective device. The activity levels for the evaluation context of the studies were adapted from [345], ranging from activity level 1 (resting state) to activity level 3 (high effort). A large share of the studies did not report standardized performance scores, but reported proof-of-concept evaluations (e.g. [498; 249; 498]).

Table A.1.: Heart rate obtained from the ear according to different sensing principles and activity levels (ALVL: 1 = resting pulse expected, 1* = resting activity with head movement, 2 = mild effort activity, 3 = high effort activity). PPG = Photoplethysmography, ECG = Electrocardiography, Mic. = Microphone, Acc. = Accelerometer, Cap. = Capactior, ITS = infrared thermophile sensor. SD = Standard Deviation, MD = mean value of difference difference in means, RMSE = root mean square error, MAE = mean absolute error, ME = mean error, fHRS = HR spectrum fidelity index, ER = Error Rate, SNR = signal-to-noise-ratio, MAPE = error as a percentage of overall mean, PAT = pulse arrival time, PTT = pulse transit time, PPV = positive predictive value, IPC = impedance cardiogram.

Sensor	**ALVL**	**Context**	**Location**	**Performance**	**Ref. Device**	**N**	**Ref.**
PPG	1	resting	lobe	ER=0.6%	ECG	13	[422]
PPG	1	resting	posterior	$f_HRS = 35 \pm 15$	PPG	10	[481]
PPG	1	resting	ear canal	ME=1.04 bpm	ECG	11	[372]
PPG	1	resting	ear canal	ME= 0.66 bpm	PPG	12	[137]
PPG	1	resting	concha	ME=-0.19 bpm	PPG	12	[137]
PPG	1	resting	posterior	ME=-0.28 bpm	PPG	12	[137]
PPG	1	sleeping	canal + lobe	show principle	ECG	6	[469]
PPG	1	sleeping	ear canal	MAE=7.01 ms	ECG	7	[69]
PPG	1	sleeping	concha	show principle	-	3-5	[439]
PPG	1	sitting	concha	MD=-0.03 bpm, 95%CI = [-2.94 and 2.88 bpm]	ECG	28	[459]
PPG	1	sitting	mastoid	MD=-0.25 bpm, SD=0.5 bpm, RMSE=1.3 bpm	PPG	10	[105]
PPG	2	standing	lobe	MD±SD= 0.62%±4.51%, r= .97	ECG	14	[382]
PPG	1	music listening	concha	0.63%<ME<0.56%	PPG	4	[381]

Continued on next page

Table A.1 – continued from previous page

Sensor	ALVL	Context	Location	Performance	Ref. Device	N	Ref.
PPG	1*	speaking	ear canal	ME=12.52%	PPG	12	[137]
PPG	1*	speaking	concha	ME=58.29%	PPG	12	[137]
PPG	1*	speaking	posterior	ME=62.60%	PPG	12	[137]
PPG	1*	chewing	ear canal	show principle	PPG	4	[475]
PPG	1*	resting+running	lobe	Max. absolute deviation = 2.5%	PPG	1	[249]
PPG	1*	running+walking	concha	RMSE=5 bpm	ECG	1	[90]
PPG	2	walking	ear canal	ME=27.14%	PPG	12	[137]
PPG	2	walking	concha	ME=58.66%	PPG	12	[137]
PPG	2	walking	posterior	ME=105.98%	PPG	12	[137]
PPG	2	walking	posterior	ME=105.98%	PPG	12	[137]
PPG	2	walking	ear canal	ME=2.77 bpm	ECG	11	[372]
PPG	2	walking	ear canal	show principle	PPG	4	[475]
PPG	2	walking	ear canal	show principle	-	2	[460]
PPG	2	walking	lobe	ER=1.7%	ECG	13	[422]
PPG	2	walking	lobe	MD±SD=-0.49%±8.65%, r= .82	ECG	14	[382]
PPG	2	walking	lobe	Max. absolute deviation = 2 bpm	PPG	6	[286]
PPG	2	walking	posterior	$f_H RS = 25 \pm 10$	PPG	10	[481]
PPG	2	walking	unspecified	show principle	ECG	1	[498]
PPG	2	moving freely	ear canal	ME=0.031±0.717 bpm	ECG	16	[53]
PPG	2	moving freely	ear canal	0.21 % false positives , 1.23 % false negatives	ECG	16	[53]
PPG	2	moving freely	ear canal	r= .87	ECG	16	[53]
PPG	2	moving freely	ear canal	MAE=8.23 ms	ECG	7	[69]

Continued on next page

Table A.1 – continued from previous page

Sensor	ALVL	Context	Location	Performance	Ref. Device	N	Ref.
PPG	3	jogging	lobe	ER=0.7%	ECG	13	[422]
PPG	3	running	ear canal	ME=29.84%	PPG	12	[137]
PPG	3	running	concha	ME=44.00%	PPG	12	[137]
PPG	3	running	posterior	ME=56.62%	PPG	12	[137]
PPG	3	running	concha	ME=-0.003 bpm, SD=2.84 bpm	ECG	1	[461]
PPG	3	running	lobe	ER=5.7%	ECG	13	[422]
PPG	3	running	lobe	MD±SD=-0.32%±10.63%, r= .76	ECG	14	[382]
PPG	3	treadmill	canal + concha	MAE=1.5 ± SD = 1.8 bpm, MAPE = 2.5%	ECG	20	[367]
PPG	other	stress test	ear canal	MD=-1.67 bpm, 14.87 bpm<SD<24.0 7bpm	PPG	3	[283]
PPG	other	stress test	concha	improving accuracy of 11.11 to 17.2 %	-	20	[169]
PPG	other	Valsalva test	lobe	PAT: 0.15<r<0.21; PTT: 0.14<r<0.31	IPC+ECG	14	[417]
PPG	other	handgrip exercise	lobe	PAT: 0.15<r<0.29; PTT: 0.16<r<0.33	IPC+ECG	14	[417]
PPG	other	thermal stress	concha	show principle	PPG	6	[468]
PPG	other	simulated altitudes	posterior	RMSE=2.61%, r= .96	PPG	12	[71]
PPG	-	unspecified	ear canal	SD=1.2 bpm, R^2>0.99	ECG	26	[470]
ECG	1	sitting	concha	R-, T-, P-, & S-waves: 10.92<SNR<17.27	PPG	1	[294]
ECG	1	sitting	mastoid	HR-waves: ME=1.3 bpm, 2.9<SNR<12.4	PPG	2	[209]
ECG	2	treatmill	mastoid	QRS: ME=2.7 ms, SD=7.8 ms	PPG	8	[502]
ECG	-	unspecified	ear canal	R-wave: SD=0.94 ms, 0.59 ms delay	PPG	10	[173]
ECG	-	unspecified	ear canal	P-, QRS-, and T-waves: r= .96	PPG	5	[477]
ECG	-	unspecified	posterior	P-, QRS-, and T-waves: show principle	PPG	1	[89]

Continued on next page

Table A.1 – continued from previous page

Sensor	ALVL	Context	Location	Performance	Ref. Device	N	Ref.
ECG	-	unspecified	beyond auricle	percentage RMS deviation=2.2-3.3%	ECG	6	[419]
Mic.	1	sitting	unspecified	1.37%<ER<1.64%	PPG	1	[128]
Mic.	1	sleeping	ear canal	r= .99	ECG	4	[160]
Mic.	1	sitting	ear canal	ME=5.6% (MAE=4.3 bpm), SE=51.2% (2.2 bpm)	ECG	25	[296]
Mic.	1	lying, sitting	concha	r= .85	ECG	37	[345]
Mic.	1	music listening	unspecified	1.42%<ER<2.42%	PPG	1	[128]
Mic.	2	treadmill	concha	r= .84	ECG	37	[345]
Mic.	3	jogging	concha	r= .75	ECG	37	[345]
Mic.	other	breathing	ear canal	MAE=4.3 bpm (2.7 cpm)	ECG	25	[296]
Acc	1	sitting	beyond auricle	show principle	PPG+ECG	8	[355]
Acc	other	Valsalva test	posterior	R^2= .66 for stroke volume	ECG	13	[184]
Acc	other	Valsalva test	posterior	R^2= .96 for R- & J-waves	ECG	13	[184]
Piezo	1	sitting	ear canal	MAD=0.62, ER=0.68%	ECG	58	[362]
Cap.	other	Valsalva test	posterior	show principle	ECG	1	[183]
ITS	-	unspecified	concha	ER=±10.5 bpm for 70% of the samples	ECG	5	[113]

A.2 Blood Oxygen Saturation

Measuring blood oxygen saturation with earables is summarized in subsubsection 3.1.1.2. In that regard, appendix Table A.2 introduces the sensor, attachment location of the sensor, reported performance, range of ground-truth blood oxygen saturation captured by the study dataset, and the number of study participants. Compared to traditional PPG, which applies LEDs to emit light, VCSEL measures the reflected light of vertical-cavity surface-emitting lasers.

Table A.2.: Performance of earable prototypes measuring the fraction of oxygenated hemoglobin in blood by measuring peripheral blood oxygen saturation.

Target	**Sensor**	**Location**	**Performance**	**Evaluation Range**	**N**	**Ref.**
SpO_2	PPG	earlobe	1% mean absolute error	93% - 98%	6	[286]
SpO_2	VCSEL	earlobe	$\approx$ 6% mean absolute error	$\approx$ 93% - 98%	1	[249]
SpO_2	PPG	posterior	2.61% root-mean squared error	70 - 100% (SaO_2)	12	[71]
SpO_2	PPG	ear canal	-0.22$\pm$1.50% mean error	96.38% - 100%	16	[53]
SpO_2	PPG	ear canal	show principle	unspecified	4	[475]
SpO_2	PPG	tragus	show principle	unspecified	11	[469]
SpO_2	PPG	ear canal	< 2% mean absolute error	96% - 100%	12	[137]
SpO_2	PPG	posterior	< 2% mean absolute error	96% - 100%	12	[137]
SpO_2	PPG	concha	< 2% mean absolute error	96% - 100%	12	[137]

A.3 Blood Pressure

Measuring blood pressure with earables is summarized in subsubsection 3.1.1.3. In appendix Table A.3 the underlying principle to measure blood pressure, the metric used for evaluation, the achieved performance of systolic and diastolic blood pressure, and the number of study participants are listed.

Table A.3.: Blood Pressure using different principles. *ECG is attached on the user's chest. SBP = systolic blood pressure, DBP = diastolic blood pressure.

Principle	**Sensor**	**Metric**	**SBP Performance**	**DBP Performance**	**N**	**Ref.**
Pulse Transit Time	PPG + ECG*	correlation	r=0.25 ± 0.30	0.33 ± 0.32	14	[417]
Pulse Transit Time	PPG + ECG*	-	show principle	unspecified	1	[168]
Pulse Amplitude	PPG + balloon	mean error	1.8 ± 7.2 mmHg	-3.1 ± 7.9 mmHg	35	[76]

A.4 Respiration

Sensing the respiration rate of a user is summarized in section 3.2. The appendix Table A.4 lists earable respiration based on the sensor applied, the measure obtained, the performance achieved, the number of study participants recruited, and the range in which ground truth respiration data was collected.

Table A.4.: Respiration rate and ins-/expiration phases measured by different sensors at the ear. MAE = mean absolute error, ME = mean error.

Sensor	**Measure**	**Performance**	**N**	**Range** ($\approx$)	**Ref.**
accelerometer	respiration rate	2.62 cycles per minute MAE	12	8 - 22 CPM	[450]
gyroscope	respiration rate	2.55 cycles per minute MAE	12	8 - 22 CPM	[450]
accel. / gyro.	respiration rate	1.64 cycles per minute ME	30	no reported	[4]
microphone	respiration rate	2.7 cycles per minute MAE	20	6 - 25 CPM	[296]
PPG	respiration rate	-0.558 $\pm$ 1.406 cycles per minute ME	16	7 - 28 CPM	[53]
PPG	respiration rate	$\approx$ 3 cycles per minute error	12	no reported	[137]
PPG	ins- / expiration	81.5% sensitivity, 86% specificity	36	4 - 21 CPM	[470]

A.5 Sleep

Sleep tracking based on earables sensing is summarised in subsubsection 3.1.2.2. In Table A.5, tracking of different sleep parameters with earable sensors is summarised. It lists sensor positions and electrode styles. Additionally, the performance and number of study participants are described for every paper.

Table A.5.: Sleep tracking based on ear-worn brain activity sensing allows to predict different sleep stages from EEG data.

Parameter	Sensor	Position	Style	Electrode	Performance	N	Ref.
sleep stages	EEG	ear canal	generic	wet	74.1% accuracy	22	[337]
sleep stages	EEG	ear canal	generic	wet	95% accuracy	8	[342]
sleep stages	EEG	ear canal	generic	wet	show principle	4	[160]
sleep stages	EEG	concha + canal	custom	wet	$\approx$ 66% accuracy	9	[315]
sleep stages	EEG	concha + canal	custom	unspecified	85.7% accuracy	1	[432]
sleep stages	EEG	periauricular	generic	wet	show principle	1	[57]
sleep spindles	EEG	concha + canal	custom	dry	show principle	12	[316]
sleep spindles	EEG	ear canal	generic	dry / wet	show principle	1	[293]
sleep latency	EEG	ear canal	generic	wet	$<$ 3 mins error for 84%	23	[10]
circadian rhythm	Infr. Ther.	ear canal	generic	-	show principle	1	[61]

A.6 Brain Activity

Brain activity sensing with earables is summarized in subsubsection 3.1.2.1. The appendix Table A.6 lists different locations evaluated with dry and wet electrodes with generic and custom fit ear-EEG devices (see subsubsection 3.1.2.1).

Table A.6.: EEG paradigms evaluated at different locations on the ear. ✓= paradigm feasible, ✗ = paradigm not feasible.

Location	Fit	Electrode	P300	MMN	ASA	ASSR	P300	SSVEP	AAR	Ref.
posterior	generic	wet	-	-	-	-	✓	✓	-	[348]
posterior	generic	wet	-	-	-	-	-	✓	-	[295]
periauricular	generic	dry	-	-	-	✓	✓	✓	✓	[166]
periauricular	generic	dry	-	-	-	✓	-	-	-	[52]
periauricular	generic	wet	✓	-	-	-	-	-	✓	[57]
periauricular	generic	wet	✓	-	-	-	-	-	✓	[115]
periauricular	generic	wet	✓	-	-	-	-	-	-	[463]
periauricular	generic	wet	-	-	-	-	-	-	✓	[489]
periauricular	generic	wet	-	-	✓	-	-	-	-	[59]
concha	generic	dry	-	-	-	✓	-	-	-	[52]
concha	custom	wet	-	-	-	-	✓	✓	-	[58]
concha + canal	generic	dry	-	-	-	✓	-	-	✓	[233]
concha + canal	custom	dry	-	-	-	✓	✓	✓	✓	[278]
concha + canal	custom	dry	-	-	-	-	-	✓	-	[279]
concha + canal	custom	dry	-	✓	-	✓	-	✓	✓	[228]
concha + canal	custom	dry	-	-	-	✓	-	-	-	[239]
concha + canal	custom	dry	✗	-	-	✓	-	✓	✓	[225]

Continued on next page

Continued from previous page

concha + canal	custom	dry	-	-	-	-	-	-	✓	[277]
concha + canal	custom	dry	-	-	-	✓	-	-	✓	[234]
concha + canal	custom	dry	-	-	-	✓	-	✓	-	[224]
concha + canal	custom	dry/wet	-	-	-	-	-	-	-	[223]
concha + canal	custom	wet	✓	-	-	✓	✓	✓	-	[240]
concha + canal	custom	wet	-	✓	-	✓	-	-	-	[314]
concha + canal	custom	dry/wet	-	-	-	-	-	-	-	[370]
concha + canal	custom	wet	✓	✓	-	-	-	-	-	[238]
concha + canal	custom	wet	-	-	-	-	-	✓	-	[482]
concha + canal	custom	wet	-	-	-	✓	-	-	-	[227]
concha + canal	custom	wet	-	-	-	✓	-	-	-	[226]
concha + canal	custom	wet	-	-	-	✓	-	-	-	[45]
concha + canal	custom	wet	✓	-	-	✓	-	-	-	[463]
ear canal	generic	dry	✓	-	-	-	-	✓	✓	[261]
ear canal	generic	dry	-	✓	-	✓	-	✓	✓	[228]
ear canal	generic	dry	✓	-	-	✓	-	✓	✓	[266]
ear canal	generic	dry	✓	-	-	-	-	-	✓	[222]
ear canal	generic	dry	-	-	-	✓	-	-	✓	[125]
ear canal	generic	wet	-	-	-	✓	-	✓	-	[342]
ear canal	generic	wet	-	-	-	✓	✓	✓	✓	[160]
ear canal	generic	wet	-	-	-	✓	✓	✓	-	[159]
ear canal	generic	wet	-	-	-	✓	-	-	-	[239]
ear canal	generic	wet	-	-	-	-	-	✓	-	[5]
ear canal	custom	dry	-	-	✓	-	-	-	-	[139]
ear canal	custom	dry	-	-	-	-	-	-	✓	[214]

A.7 Drowsiness

Drowsiness detection based on earables sensing is summarised in subsubsection 3.1.2.3. In Table A.7, tracking of drowsiness with ear EEG is summarised. It lists sensor positions and electrode styles. Additionally, the performance and number of study participants are described.

Table A.7.: Drowsiness detection based on ear-worn brain activity sensing EEG data.

Parameter	**Sensor**	**Position**	**Style**	**Electrode**	**Performance**	**N**	**Ref.**
alpha power increase	EEG	periauricular	generic	wet	show principle	10	[57]
microsleep	EEG	periauricular	generic	wet	87% precision	19	[374]
drowsiness	EEG	ear canal	generic	wet	80% accuracy	23	[335]

A.8 Emotions

Emotion tracking with earables is described in subsubsection 3.1.4.1. Table A.8 summarizes earable works that looked into tracking emotions of the wearer according to the applied scale (dimensional or categorical), the specific emotions predicted, the sensor used, the performance achieved, and the number of study participants recruited.

Table A.8.: Emotions were classified from the ear from audio data, brain activity and indirectly from body motion intensity.

Scale	**Emotions**	**Sensor**	**Performance**	**N**	**Ref.**
dimensional	valence (high vs. low)	in-ear EEG	71.07% accuracy	13	[35]
dimensional	valence (high vs. low)	in-ear EEG	94.1% accuracy	12	[269]
dimensional	valence (high vs. low)	accelerometer	show principle	3	[146]
dimensional	arousal (high vs. low)	in-ear EEG	72.89% accuracy	13	[35]
categorical	happiness, sadness, calmness, fear	in-ear EEG	53.72% accuracy	13	[35]
categorical	excited, relaxed, negative	in-ear EEG	58.8% accuracy	12	[269]
categorical	neutral, upset, happy, angry	microphone	91% F1-score	24	[229]

A.9 Eating

In subsection 4.1.4 tracking of eating-related events with earables is summarized. In Table A.9 below, earable works that have addressed tracking eating-related activities are summarized according to the parameter they aimed to predict, the study setup used (e.g., field, lab), the food ingested (e.g., free, N pre-defined foods), the sensor applied, the location where the sensor was placed, the achieved performance, and the number of study participants.

Table A.9.: Tracking of eating related events with earables. * = an additional sensor beside the sensor on the ear was used, ** = manual inter rater agreement (ICC)

Parameter	Setup	Ingest	Sensor	Location	Performance	N	Ref.
bite events	lab	4 foods	piezoelectric + bone mic.	inferior + mastoid	ICC=0.935**	21	[414]
chew events	lab	free	2 × microphone	canal + lateral	72% precision	6	[310]
chew events	lab	5 foods	microphone	ear canal	98.07% accuracy	1	[346]
chew events	lab	3 foods	microphone	ear canal	>60% precision	8	[14]
chew events	lab	4 foods	piezoelectric + bone mic.	inferior + mastoid	ICC=0.928**	21	[414]
chew strength	lab	3 foods	pressure sensor	ear canal	show principle	15	[200]
chew strength	lab	3 foods	microphone	ear canal	>59.0% accuracy	8	[14]
swallow events	lab	4 foods	piezoelectric + bone mic.	inferior + mastoid	ICC=0.943**	21	[414]
drinking detection	lab	1 drink	microphone + bone mic.	canal + lateral	56.6% accuracy	5	[424]
drinking detection	lab	1 drink	2 × microphone	canal + lateral	30% recall	51	[366]
predict food type	lab	8 foods	2 × microphone	canal + lateral	79% accuracy	51	[366]
predict food type	lab	4 foods	microphone	ear canal	86.6% accuracy	2	[13]
predict food type	lab	3 foods	microphone	ear canal	94.0% accuracy	8	[14]
intake progress	field	1 meal	camera	lateral	show principle	1	[273]

Continued on next page

Continued from previous page

eating detection	field	free	bone microphone	mastoid	87.9% accuracy	14	[56]
eating detection	field	free	microphone	lateral	76.8% accuracy	4	[150]
eating detection	semi	33 foods	microphone	ear canal	88.0% accuracy	22	[360]
eating detection	lab	12 foods	microphone	lateral	95.5% accuracy	28	[150]
eating detection	lab	7 foods	2 × microphone	canal + lateral	88.5% accuracy	51	[366]
eating detection	lab	6 foods	bone microphone	mastoid	90.0% accuracy	20	[55]
eating detection	lab	5 foods	microphone	lateral	94.0% accuracy	5	[281]
eating detection	lab	7 foods	2 × microphone	canal + lateral	77.6% accuracy	40	[364]
eating detection	lab	5 foods	acc. + gyr.	lateral	95.0% accuracy	5	[281]
eating detection	lab	unknown	acc. + gyr.	lateral	84% accuracy	20	[256]
eating detection	field	free	acc. + gyr. + mag.*	posterior	93.0% accuracy	10	[49]
eating detection	lab	free	acc. + gyr. + mag.	posterior	86.0% accuracy	16	[49]
eating detection	lab	free	acc. + gyr. + mag. + prox.	posterior + canal	81.5% accuracy	16	[49]
eating detection	lab	5 foods	acc. + gyr. + microph.	lateral	97.0% accuracy	5	[281]
eating detection	field	free	proximity sensor	ear canal	82.3% accuracy	6	[48]
eating detection	lab	3 foods	proximity sensor	ear canal	90.4% accuracy	20	[48]
eating detection	lab	free	proximity sensor	ear canal	71.5% accuracy	16	[49]
eating detection	field	free	piezoelectric*	inferior	89.8% accuracy	12	[140]
eating detection	lab	5 foods	piezoelectric	inferior	81.0% accuracy	20	[415]
eating detection	semi	33 foods	PPG	ear canal	75.3% accuracy	22	[360]
eating detection	lab	6 foods	EMG	Mastoid	84.0% accuracy	20	[55]
eating detection	lab	6 foods	EMG + microphone	Mastoid	94.0% accuracy	20	[55]

A.10 Gait

Gait tracking with earables is summarized in subsection 4.1.3. Commonly, gait parameters and related events are predicted by acceleration forces measured at the ear. Table A.10 summarizes gait tracking earable works.

Table A.10.: Gait-related parameters that were looked into with earables.

Gait-Related Parameter	**Performance**	**N**	**Ref.**
timing of heel contact, and toe off events	35.38±3.22, 73.05±7.24 ms error	10	[210]
timing of swing, stance, and stride time	35.5±4.0, 36.9±3.8, 17.9±2.3 ms error	10	[211]
timing of step cycle and asymmetry	20±50, 10±70 ms mean difference	64	[32]
timing of step cycle	30 ms mean difference	43	[30]
detect steps (very slow, slow, normal, run)	95% accuracy	3	[387]
detect truncal impairment	> 97% accuracy	22	[31]
detect lower limb impairment	95% accuracy	20	[25]
detect freeze of gait	95.6% accuracy	16	[280]
detect simulated lower limb impairment	> 97% accuracy	10	[31]
Timed Up-and-Go (TUG) fall risk [308]	elderly at risk correctly identified	28	[244]
post-surgery gait pattern regularity	regluarity improves 24 weeks post-op	17	[27]
post-surgery asymmetry, stride and step time	metrics improve over a year post-op	16	[213]
Knee Injury & Osteoarthritis Outcome Score [402]	r = -0.68	12	[210]
Tinetti Gait and Balance Assessment (TGBA) [465]	360° turn, stand w/ eyes closed feasible	28	[244]
Parkinson's-related tremor noise removal	noise can be separated from gait	9	[365]

A.11 Ear and Earable Input

Interaction on the ear and earable input is summarized in subsection 5.1.1. The Table A.11 below summarizes how manipulation can be performed at varying locations based on different sensing principles.

Table A.11.: Summary of works that looked into manipulation of the ear and earable.

Manipulation	Location	Sensor	Ref
touch	earable	capacitive	[79]
touch	earable	capacitive	[272]
touch	earable	capacitive and resistive	[486]
touch	ear	camera	[259]
taps	earable	modified speaker	[290]
taps and sliding gestures	ear; cheek; temple; mandible	microphone	[494]
ear deformation	ear	proximity sensors	[241]
tensor tympani contraction	ear canal	pressure sensor	[453]
wearable state	earable	capacitive	[79]
wearable state	earable	proximity sensors	[302]
wearable state	earable	microphone + emitted ultrasound	[257]
3D position	earable	acoustic motion tracking	[85]

A.12 Hand Gestures and Location

Hand gestures and hand location tracking are described in subsection 5.1.2. The Table A.12 below describes where the interaction takes place, how activation is performed, what sensing principle is used, the achieved performance, and the number of study participants.

Table A.12.: Hand tracking embedded inside earables enables interaction away from the ear.

Location	**Activation**	**Sensor**	**Performance**	**N**	**Ref.**
mid-air at mouth	obstruct mouth by hand	microphone + emitted sound	96.5% F1-score	12	[497]
mid-air at earable	swipe 1 to 4 fingers	proximity sensor	99% accuracy	2	[312]
mid-air at earable	42 sign language words	microphone + emitted sound	93.8% accuracy	8	[215]
mid-air	hand gestures	camera + mini-projector	show principle	1	[443]

A.13 Mouth-Based Interaction

In subsection 5.1.4 mouth-based interaction is summarized. In the Table A.13 below, mout-related gestures based on jaw, teeth, and tongue are introduced according to the type of gestures, sensors, achieved accuracy, and number of study participants.

Table A.13.: Movement of the jaw, tongue, and teeth was applied to offer subtle interaction to the user.

Mouth-Related Gesture	**Sensor**	**Accuracy**	**N**	**Ref.**
6 jaw movement gestures	pressure sensor	90.7%	12	[16]
4 open mouth levels	pressure sensor	87.5%	12	[16]
10 target points at teeth by tongue	EMG, EEG, capacitive sensing	88.61%	15	[344]
6 target points at teeth by tongue	microphone + emitted sound	94.8%	20	[86]
point tongue left, front, right	pressure sensor	66 - 96% (per user)	5	[287]
press tongue against roof of mouth	photo reflective	> 99%	5	[444]
click 5 different pairs of teeth	microphone	78%	20	[22]
6 tap and slide teeth gestures	modified speaker	> 90%	18	[389]
13 teeth tap gestures	passive microphone + gyro.	90.9%	11	[435]

A.14 Facial Expressions

In subsection 5.1.5 facial expression tracking with earables is described. In Table A.14 facial expression tracking works with earables are summarized based on the types of facial expressions detected, sensors applied, performance achieved, and number of study participants recruited.

Table A.14.: Facial expressions were explored for discreet interaction with many works selecting the ideal subset of expressions for increased classification performance. * = final set selected from a bigger data collection

Facial Expressions	**Sensor**	**Performance**	**N**	**Ref.**
eye wink, head right, open mouth, say sh, smile*	EMG (EarFS)	90% precision	3	[304]
face up/down, tilt head left/right, jaw left/right*	mic. + emitted sound	90% F1-Score	11	[11]
smile, frown*	Acclerometer	85% F1-Score	9	[264]
15 facial expressions	Miniature Cameras	88.6% accuracy	9	[96]
32 facial expressions	Accel. + Gyro.	89.9% accuracy	12	[472]

A.15 Shape-Based Authentication and Identification

Authentication and identification as described in section 7.4. In Table A.15 different works are summarized. Verification performance reported as false rejection rate (FRR), false acceptance rate (FAR), and equal error rate (ERR). Equal error rate is the decision boundary at which FAR and FRR are equal. For identification, the rate of correct identifications is reported.

Table A.15.: Perfomance of different earable verification and identification methods. * indicates that no FAR/FRR combination was highlighted (FRR = FAR = EER).

		Verification			**Identificat.**		
Target	**Probe Signal**	*FRR*	FAR	*EER*	*Rate*	**N**	**Ref.**
ear	audible chirp (1.5 - 22 kHz)	*	*	0.8%	-	31	[9]
ear	audible seq. (6 Hz - 12 kHz)	*	*	14.9%	-	50	[118]
ear	audible chirp (100 Hz - 10 kHz)	*	*	$\approx$ 5%	-	27	[118]
ear canal	audible chirp (20 Hz - 6 kHz)	2.6%	2.2%	-	-	20	[151]
ear canal	audible chirp (1.5 - 22 kHz)	*	*	1.0%	-	31	[9]
ear canal	audible chirp (1 Hz - 18 kHz)	*	*	0.28%	-	25	[288]
ear canal	audible seq. (200 Hz - 16 kHz)	*	*	0.97%	-	45	[21]
ear canal	inaudible chirp (18 - 48 kHz)	*	*	< 0.01%	-	25	[288]
ear canal change	inaudible chirp (16 - 23 kHz)	*	*	$\approx$ 4%	-	24	[483]
cochlea response	audible click stimulus	1.55%	0.0%	0.02%	99.44%	54	[276]
skull	speaking-induced body sound	*	*	3.64%	-	23	[152]
skull	voice 'EMM' sound	*	*	1.28%	-	34	[274]
body (arm to ear)	laptop leakage current	< 8%	< 10%	-	93.6%	15	[121]