Towards Detecting Alzheimer’s Disease with Central Auditory and Physiological Measures

by

Miriam BOUTROS

MANUSCRIPT-BASED THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT OF A MASTER’S DEGREE WITH THESIS
M.A.Sc.

MONTREAL, DECEMBER 8, 2025

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Prof. Rachel Bouserhal, Thesis supervisor
Department of Electrical Engineering, École de technologie supérieure

Prof. Christopher Niemczak, Thesis Co-Supervisor
Geisel School of Medicine, Dartmouth College

Prof. Claude Thibeault, Chair, Board of Examiners
Department of Electrical Engineering, École de technologie supérieure

Prof. Sylvie Ratté, Member of the Jury
Department of Software Engineering and IT, École de technologie supérieure

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND THE PUBLIC

ON NOVEMBER 24, 2025

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

First, I am deeply grateful for this journey, which has allowed me to learn and grow immensely. I sincerely thank my supervisors, Rachel Bouserhal and Chris Niemczak, for their invaluable guidance throughout this research. I am especially thankful to Rachel for fostering a nurturing environment where expressing ideas and asking questions was always encouraged. I also appreciate Chris for his consistent support and guidance, even from afar in the United States.

I am grateful to the Mitacs internships, which provided me with the skills necessary to complete this work. I would also like to thank the participants who generously gave their time and were eager to contribute to this research.

To my family: my mother, for being my emotional support and safe haven throughout the years; my father, for his constant encouragement; my sister, for being my voice of reason and someone I look up to; and my aunt, for facilitating my move to Canada and helping me adapt. I am also deeply thankful to my partner for his immense support, as well as to my friends in Lebanon and Canada for cheering me on every step of the way.

Vers la détection de la maladie d'Alzheimer à l'aide de mesures auditives centrales et physiologiques

Miriam BOUTROS

RÉSUMÉ

La maladie d'Alzheimer (MA) est la principale cause de démence, et le nombre de cas devrait augmenter considérablement avec le vieillissement de la population mondiale. Le déclin cognitif associé à la MA débute souvent par des symptômes subtils, tels que des difficultés à comprendre la parole dans des environnements bruyants, avant d'évoluer vers des troubles de la mémoire et des déficiences sévères. Un diagnostic précoce est essentiel, car il permet une prise en charge clinique rapide et ouvre la voie à des changements de mode de vie susceptibles de retarder ou de prévenir la démence. Cependant, les méthodes diagnostiques actuelles, incluant les évaluations neuropsychologiques et la neuroimagerie, demeurent coûteuses et inadaptées au dépistage à grande échelle. Cela souligne la nécessité de développer des approches alternatives précises, accessibles et non invasives, pouvant compléter les méthodes diagnostiques traditionnelles.

Des recherches récentes mettent en évidence les fonctions auditives centrales, langagières et physiologiques comme des biomarqueurs prometteurs pour la détection précoce de la MA. Les déficits du traitement auditif central, tels qu'une diminution de la capacité à comprendre la parole dans le bruit, sont liés à des régions corticales affectées par la maladie et peuvent être évalués au moyen de tests simples et non invasifs. De même, les altérations du discours et du lexique, ainsi que certains bio-signaux tels que la fréquence cardiaque et le diamètre pupillaire, peuvent offrir des indices diagnostiques complémentaires. Les dispositifs « hearables » constituent une plateforme pratique pour la capture de ces signaux multimodaux à l'aide de microphones intra-auriculaires, grâce à l'effet d'occlusion qui amplifie les signaux de basse fréquence lorsque l'oreille est obstruée. Toutefois, il existe encore un manque de bases de données multimodales collectées auprès de personnes atteintes de MA ou de trouble cognitif léger (MCI) à l'aide d'un hearable.

Ce mémoire comble cette lacune en présentant la base de données GARDENIA (Gaze and Auditory Response Database for Evaluating Neurocognitive Impairment and Alzheimer's disease), un ensemble de données multimodal recueilli auprès de 20 participants, incluant des individus atteints de MA, de MCI et des témoins cognitivement sains. La base de données comprend des tests de traitement auditif central, une tâche de description d'image, des biosignaux intra-auriculaires et des données de suivi oculaire. Les résultats indiquent que les participants cognitivement atteints ont obtenu de moins bonnes performances que les témoins à l'ensemble des tests auditifs centraux, les tâches dichotiques présentant la plus forte valeur prédictive. Ils démontrent également la capacité des hearables à administrer efficacement les tests auditifs centraux. Cependant, l'analyse des battements cardiaques extraits des signaux intra-auriculaires à l'aide de Tempbeat a mis en évidence la nécessité de développer des algorithmes d'extraction de biosignaux robustes, insensibles aux mouvements de la mâchoire. Dans l'ensemble, ce travail

présente GARDENIA comme une ressource précieuse pour la recherche et le développement d'outils de traitement des biosignaux visant à améliorer la détection de la maladie d'Alzheimer.

Mots-clés: maladie d'Alzheimer, hearables, tests de traitement auditif central, signaux biologiques

Towards Detecting Alzheimer's Disease with Central Auditory and Physiological Measures

Miriam BOUTROS

ABSTRACT

Alzheimer's disease (AD) is the leading cause of dementia, with cases expected to rise substantially as the aging global population. Cognitive decline in AD often begins with subtle symptoms, such as difficulty understanding speech in noisy environments, before advancing to memory loss and severe impairments. Early diagnosis is critical, as it enables timely clinical care and creates opportunities for lifestyle changes that may delay or prevent dementia. However, current diagnostic methods, including neuropsychological assessments and neuroimaging, are expensive and unsuitable for large-scale screening, emphasizing the need for alternative accurate, accessible, and non-invasive approaches that could complement the traditional diagnostic methods.

Recent research highlights central auditory, speech, and physiological functions as promising biomarkers for early detection of AD. Central auditory processing deficits, such as reduced ability to understand speech in noise, are linked to cortical regions affected by AD and can be evaluated using accessible and non-invasive tests. Similarly, speech and lexical changes along with biosignals, such as heart rate and pupil diameter, may provide potential additional diagnostic insights. Hearable devices offer a practical platform for capturing these multimodal signals using in-ear microphones due to the occlusion effect that amplifies low-frequency signals when the ear is occluded. However, there is a lack of multimodal datasets collected from individuals with AD and mild cognitive impairment (MCI) using a hearable.

This thesis addresses this gap by introducing the Gaze and Auditory Response Database for Evaluating Neurocognitive Impairment and Alzheimer's disease (GARDENIA), a multimodal dataset collected from 20 participants, including individuals with AD, MCI, and cognitively unimpaired controls. The dataset contains central auditory processing tests, a picture description task, in-ear biosignals, and eye-tracking data. Results indicate that cognitively impaired participants performed worse than controls across all central auditory tests, with dichotic tasks showing the strongest predictive value. Also, they demonstrate the capability of hearables to successfully administer central auditory tests. However, the analysis of heartbeats extracted from in-ear signals using Tempbeat highlighted the need for robust bio-signal extraction algorithms that are not affected by jaw movements. Overall, this work presents GARDENIA as a valuable resource for researching and developing tools for bio-signal processing that can help detecting AD.

Keywords: Alzheimer's disease, hearables, central auditory tests, biosignals

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Alzheimer’s Disease
BoAW	Bag of Audio Words
CAT	Central Auditory Test
CDR	Clinical Dementia Rating
CI	Cognitively Impaired
CSF	Cerebrospinal Fluid
CU	Cognitively Unimpaired
DDT	Dichotic Digits Test
DSI	Dichotic Sentence Identification Test
DTI	Diffusion Tensor Imaging
EMREO	Eye-Movement Related Eardrum Oscillation
GAP	Gap Detection Test
GARDENIA	Gaze and Auditory Response Database for Evaluating Neurocognitive Impairment and Alzheimer’s Disease
GSR	Galvanic Skin Response
GUI	Graphical User Interface
HF	High-Frequency
HINT	Hearing in Noise Test
HR	Heart Rate
HRV	Heart Rate Variability

HTL	Hearing Threshold Level
IEM	In-Ear Microphone
LC	Locus Coeruleus
LF	Low-Frequency
LFHF	Low-Frequency to High-Frequency (LFHF)
MAE	Mean Absolute Error
MCI	Mild Cognitive Impairment
MCL	Most Comfortable Level
MCSA	McGill University Research Centre for Studies in Aging
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
OEM	Outer-Ear Microphone
PET	Positron Emission Tomography
PNS	Parasympathetic Nervous System
PPG	Photoplethysmogram
PTA	Pure Tone Average
RMSSD	Root Mean Square of Successive Differences Between Normal Heartbeats
RRI	R-R Interval

SDNN	Standard Deviation of the Interbeat Interval of Normal Sinus Beats
SNR	Signal to Noise Ratio
SNS	Sympathetic Nervous System
SRT	Speech Reception Threshold
SSI-ICM	Synthetic Sentence Identification with Ipsilateral Competing Message
TDT	Triple Digits Test
TRIAD	Translational Biomarkers of Aging and Dementia

LIST OF SYMBOLS AND UNITS OF MEASUREMENTS

bpm	Beats per minute
dB	Decibel
dBA	A-weighted decibel
dB HL	Decibel Hearing Level
Hz	Hertz
ms	Millisecond

INTRODUCTION

0.1 Context

Alzheimer's disease (AD) is the most common form of dementia and accounts for the majority of dementia cases worldwide (Alzheimer's Disease International, 2022). In Canada, more than 600,000 individuals were living with dementia in 2022, and this number is projected to double by 2030 (Alzheimer Society of Canada, 2023). With the steady increase in life expectancy and the global aging population, the prevalence and societal burden of AD are expected to rise substantially in the coming decades. Early diagnosis of AD is crucial and preferable, not only for seeking clinical care, but also for adopting preventive proactive measures. Lifestyle changes, such as including physical activity, adopting a healthy diet, limiting smoking and alcohol, and engaging in social and cognitive activities, may delay the onset of cognitive decline, and is some cases, prevent dementia (Kivipelto, Mangialasche & Ngandu, 2018). Hence, detecting AD in its preclinical stage highlights a critical. window where lifestyle interventions might define the trajectory of the disease and its effect for the coming years. The rise of AD in the aging population and the importance of timely diagnosis highlight the pressing need for diagnostic approaches that are not only accurate, but also scalable, accessible, and non-invasive.

Traditional diagnostic practices, including neuropsychological assessments and neuroimaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans, remain the most common reliable diagnostic tools, however, they are often time-intensive, resource-demanding, and not well-suited for population-wide screening (Alzheimer's Association, 2025a). As a result, there is a critical gap in the availability of early detection tools that could complement or precede these conventional methods. Addressing this gap has become a central focus of recent research, which has expanded considerably over the past several decades. These efforts have identified a number of promising physiological, behavioral, and cognitive measures that may serve as potential biomarkers for the early detection of AD.

Research on the connection between auditory function and AD has gained increasing attention as a means to address this gap. The peripheral auditory system consists of the ear, including the cochlea, and is responsible for detecting and transmitting auditory signals to the brain. While peripheral hearing loss has long been associated with cognitive decline (Lin *et al.*, 2013), the central auditory processing system, while not mutually exclusive, has been linked to neuropathological changes of AD. The central auditory processing system is primarily the cortical aspect of the auditory system that is responsible for processing and interpreting the signals captured by the peripheral auditory system. This system employs functions related to memory, language, attention, and sound localization. These executive functions are part of the cortical areas that are affected by AD, hence, decline in central auditory processing functions might indicate the presence of AD. Tasks that require understanding speech in noise or with a competing message are normally used to assess the state of the central auditory processing system. Intensive research was conducted using these tests as predictors of AD (Gates, Anderson, McCurry, Feeney & Larson, 2011), and evidence demonstrates their potential of being an AD diagnostic tool as these tests are non-invasive and cost-effective.

Speech and lexical features have also merged as valuable tools for identifying individuals with AD (Hernández Domínguez, 2019). Subtle changes in speech patterns, such as increased pauses and repetition, are associated with early stages of AD (Forbes-McKay & Venneri, 2005). Similarly to central auditory processing functions, speech and language rely on executive functions of the brain and memory, which are affected by AD. Studying the changes in these features and their deterioration over time might provide a non-invasive tool for AD diagnosis.

Research has also delved into analyzing biosignals in order to find biomarkers for AD. For example, individuals with AD might experience higher stress levels than healthy individuals while performing a certain task (Nair *et al.*, 2023). This increase stress is equivalent to going into the sympathetic state or "fight-or-flight" mode, which is identified by an increase in heart rate, pupil diameter, and galvanic skin response (GSR). Researching and analyzing these signals provide an opportunity to better understand the subtle manifestations of AD that might be biomarkers for early AD.

To answer the need for a non-invasive, cost-effective, and scalable diagnostic tool, a technology that is able to perform central auditory tasks and capture speech and biological signals is needed. Wearable technologies offer a promising solution as they provide a continuous non-invasive monitoring of physiological signals. Among these devices, hearables, in-ear wearables, have the potential to tackle this challenge for numerous reasons. First, they can be equipped with in-ear microphones (IEMs) that can capture physiological signals from inside the ear due to the occlusion effect. The occlusion effect takes place when the ears are occluded, which amplifies low-frequency, such as heartbeats, signals propagated by tissue and bone conduction. Hence, they can cover the collection of biosignals that might be biomarkers for AD. Second, they can also be equipped by outer-ear microphones (OEMs), in addition to IEMs, which means that they can also capture speech coming from inside and outside the ear, monitoring the speech-based AD biomarkers while being able to transmit to the wearer audio information coming from the environment. Third, they might be able to administer central auditory tests (CATs) that evaluate the state of the central auditory processing system. Finally, hearables are non-invasive and accessible since they are already integrated in the daily lives of the population as noise-canceling tools or hearing aids in the aging population.

0.2 Motivation

To achieve the vision presented in the context above, a proof of concept is needed to test the feasibility of using hearables to capture the biological data needed to identify individuals with AD while testing central auditory function. Despite the growing interest in intra-aural wearables, the field is lacking the necessary data to fulfill this need. The data needs to be taken from individuals with AD or mild cognitive impairment and should capture various biosignals through both, the IEMs of the hearable as well as the respective ground truths while performing CATs.

0.3 Objectives

The main goal of this thesis is to mitigate the limited amount of available data by achieving the following goals:

1. **Adapt central auditory tests to be performed through a hearable while capturing biosignals**

   Calibrate the hearable to present the stimuli at controlled levels and developing a graphical user interface to facilitate the administration of the tests. The selected tests are speech in noise tests such as hearing in noise test (HINT) and triple digits test (TDT), and dichotic tests like dichotic digits test (DDT) and dichotic sentence identification test (DSI). Additionally, it was hypothesized that impaired participants would perform worse on CATs and hence, these tests could be used to distinguish between impaired individuals and unimpaired controls.
2. **Create a multimodal database**

   Build a comprehensive multimodal database that contains data from AD patients, individuals with mild cognitive impairment (MCI), and cognitively healthy controls. The database should contain in-ear signals, biosignals, such as heart beats and eye movements, captured using research approved sensors to serve as ground truths to the biosignals extracted from the in-ear signals, as well as the results of CATs.
3. **Validate part of the database**

   Due to the multimodality of GARDENIA and the limitation of the exploratory scope of this work, only part of the signals captured is to be analyzed within the scope of this master's thesis. heart rate (HR) and heart rate variability (HRV) are chosen to be extracted and observed. This choice is due to the availability of an algorithm that extracts heartbeats from signals captured by the IEM, and a corresponding ground truth represented by photoplethysmogram (PPG) data.

0.4 Contributions

This work resulted in the creation of a multimodal database called GARDENIA. This database was presented in a journal article, titled "A Multimodal Database for Detecting Alzheimer's Disease with Central Auditory and Physiological Measures: A Pilot Study" submitted to Hearing Research, that contains a thorough analysis of the results of the CATs. Finally, the five graphical user interfaces, that include the CATs, used in the data collection and GARDENIA are available to future use by reasonable request and as long as they adhere to the Québec laws in place.

0.5 Structure of Thesis

This thesis is structured as follows:

0.5.1 Chapter 1 - Literature review

This chapter presents the literature review that was conducted on the three main themes. In the first section, Alzheimer's disease is explained with its current standard diagnostic methods. In the second section, the chapter tackles new diagnostic tools that have the potential to complement or replace the standard ones in the early stages. This section discusses biosignals, including speech and eye data, and the auditory system with an emphasis on central auditory function and its related tests. Finally, the third section dives deeper into hearable technologies and how they were used in previous studies.

0.5.2 Chapter 2 - A Multimodal Database for Detecting Alzheimer's Disease with Central Auditory and Physiological Measures: A Pilot Study

This chapter presents the main contribution of this work, which is a journal article submitted to Hearing Research on October 2025 titled "A Multimodal Database for Detecting Alzheimer's Disease with Central Auditory and Physiological Measures: A Pilot Study". The article presents the methodology that was followed in the data collection, in addition to a brief description of the content of the database. It also presents the results of the central auditory tests as a way to show one aspect of the database and its utility.

0.5.3 Chapter 3 - Preliminary Results from Hearable Data and PPG

This chapter presents a primarily analysis of the quality of the heart rate extracted from in-ear signals using an existing in-ear heartbeat extraction algorithm. In the chapter, heartbeats extracted from the IEM are compared to those extracted by a PPG. HRV features are extracted as well in an effort to identify any patterns related to the administration of the CATs.

0.5.4 Chapter 4 - Conclusions and Future Work

The final chapter summarizes the conclusions drawn from this research, re-iterates the contributions made in Alzheimer's disease detection and wearable devices, and suggests some recommendations for future work.

CHAPTER 1

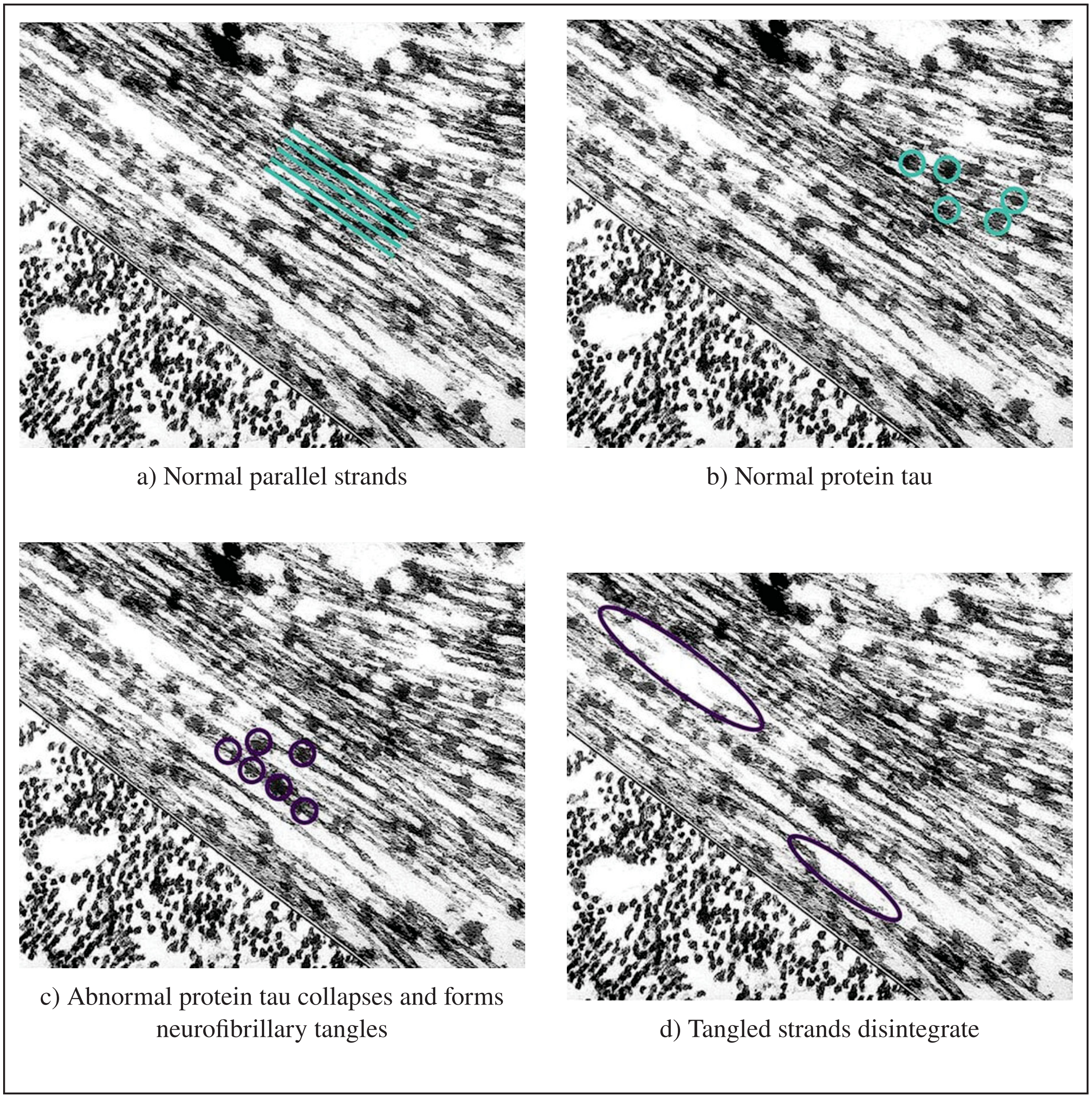
LITERATURE REVIEW

1.1 Alzheimer’s disease and its diagnosis

1.1.1 Definition

AD is a progressive neurodegenerative disorder that leads to pathological changes in the brain represented by the formation of amyloid plaques and neurofibrillary tangles (DeTure & Dickson, 2019). The human brain is made up of a network of brain cells called neurons that form brain regions responsible for various functions such as memory, attention, speech, and language (Nauta & Feirtag, 1979). Scans of patients with AD show that the brain shrinks due to the death of neurons which causes a deterioration in cognitive functions (Terry, 2006). The death of neurons is potentially due to two factors: plaques that are formed by small clumps of beta-amyloid protein, and neurofibrillary tangles, shown in Figure 1.1, made of tau protein, that prevent the required nutrients to be transported through cells (DeTure & Dickson, 2019).

There are different stages of AD ranging from mild to severe, with early signs of AD potentially manifesting 20 to 30 years before noticeable clinical symptoms appear (Tarawneh, Menegola, Peou, Tarawneh & Jayakody, 2022). During the early stages, areas of the brain responsible for thinking, planning, learning, and memory are affected. Then, during mild to moderate stages, speech and language are impacted. Finally, during the severe stages of AD, individuals experience dementia and have reduced motor control making them unable to walk or communicate verbally (Alzheimer’s Association, 2025a). According to the Alzheimer Society of Canada (2024), Alzheimer’s dementia is widely recognized as the most common form of dementia where patients are not able to communicate, care for themselves, or recognize their loved ones.



a) Normal parallel strands

b) Normal protein tau

c) Abnormal protein tau collapses and forms neurofibrillary tangles

d) Tangled strands disintegrate

Figure 1.1 Progressive changes in tau protein structure
Taken and adapted from Alzheimer's Association (2025b)

1.1.2 Traditional diagnostic methods

Detecting AD early, before symptoms appear could radically change AD treatment. Early detection could slow it's progression and help delay or even prevent dementia through lifestyle modifications such as limiting alcohol consumption, quitting smoking, managing obesity, and treating depression (Rasmussen & Langerman, 2019). The diagnosis of AD typically involves neurophysiological evaluations and brain imaging techniques such as MRI and PET scans (Weller & Budson, 2018), as AD affects primarily the medial temporal lobes, which distinguishes it from other conditions of MCI (Dickerson & Sperling, 2008). MCI is the "symptomatic predementia stage" that encompasses numerous cognitive diseases that affect cognitive areas of the brain that could be both related or not related to memory (Roberts & Knopman, 2013). MCI due to AD is one of the subtypes of MCI. It can be recognized by MRI and PET scans that show the characteristics of the effects of AD on the brain. People with MCI due to AD are more likely to develop AD (Roberts & Knopman, 2013). Other available AD diagnostic methods also include the identification of biomarkers in cerebrospinal fluid (CSF), blood tests, and diffusion tensor imaging (DTI) (Bomasang-Layno & Bronsther, 2021). In addition to imaging and biomarker analysis, standardized cognitive assessments are also used in the diagnosis of AD. The mini-mental state examination (MMSE) is a ten minute exam containing eleven questions that test orientation, registration, attention and calculation, recall, and language (Folstein, Robins & Helzer, 1983). It is scored from 0-30, with thirty indicating a cognitively healthy individual and 23 and below indicating cognitive impairment. The test has a sensitivity of 0.85 and specificity of 0.90 (Creavin *et al.*, 2016). Which means that 85% of the population is correctly classified as having dementia, while 15% is wrongly classified as having dementia. Also, 90% of the population is correctly classified as not having dementia, while the 10% is wrongly classified as not having dementia. Another interview-based test is the clinical dementia rating (CDR) that assesses the participant's memory, problem solving, orientation, and hobbies by asking the participant and their caregiver questions (Berg, 1984). CDR scores vary between zero and three, with zero indicating no dementia, one indicating MCI, two indicating moderate cognitive impairment, and three indicating severe cognitive impairment (Khan, 2016). The test has a sensitivity of 0.95 and specificity of 0.94 (Juva *et al.*, 1995). Another test that evaluates the cognitive state is the Montreal cognitive assessment (MoCA) test. MoCA is a ten-minute test that assesses memory, attention, executive function, language, recall, and orientation (Nasreddine *et al.*, 2005). Similarly to the MMSE, it is scored between 0 and 30, with 25 and lower indicating abnormal cognitive health. It has a sensitivity of 0.90 and specificity of 0.87 (Nasreddine *et al.*, 2005). However, when used with different cutoffs the sensitivity and the specificity change (Dautzenberg, Lijmer & Beekman, 2020). Additionally, educational background affects the results of MoCA where cognitively impaired individuals with more years of education score higher than predicted (Julayanont, Phillips, Chertkow & Nasreddine, 2012).

In summary, traditional diagnostic methods for AD are either imaging and biomarker-based tools, which are expensive, time-consuming, and require specialized medical equipment and trained personnel, or cognitive tests, which are more accessible but depend on interview-based evaluations. These assessments are often subjective, influenced by the clinician's judgment and the educational background of the participant. Hence, there is a need for a diagnostic approaches that are more accessible, accurate, objective, scalable, non-invasive, and capable of detecting cognitive decline at its earlier stages.

1.1.3 Other potential diagnostic methods

1.1.3.1 The auditory system

Previous studies indicate that pathological alterations in the human auditory system can be objectively assessed and may help distinguish between normal age-related cognitive decline, MCI, and AD (Swords, Nguyen, Mudar & Llano, 2018a).

The auditory system consists of the peripheral and the central auditory system. While this is not a perfect dichotomy, this distinction helps separate two fundamental contributions to Alzheimer's disease development. The peripheral auditory system includes the external, middle, and inner ear, along with the cochlea (i.e., the "hearing" organ). Its primary role is detecting sound and transmitting the auditory signals to the brain (Musiek & Baran, 2018). The pure tone average (PTA) audiometry test is used to assess peripheral hearing. During this test, pure tone stimuli are presented to the individual and the lowest hearing threshold for frequencies between 125 Hz to 8 kHz are obtained. Based on these thresholds, peripheral hearing loss can be detected and diagnosed. Although normal aging leads to hearing loss, known as presbycusis, and subsequently, cognitive changes, the influence of such hearing loss on the progression is still not fully understood (Swords *et al.*, 2018a).

On the other hand, the central auditory system is responsible for processing and interpreting the sound signals transmitted by the peripheral auditory system. The central auditory system is therefore responsible for processing of complex sounds, like speech, decoding lexical context, and even tracking the source of the signals by relying on various cells and regions within the auditory cortex, which contributes to sound interpretation and speech perception (Musiek & Baran, 2018). Individuals with central auditory processing impairment often struggle with speech perception, especially in noisy environments. This impairment can be caused by AD where amyloid plaques and neurofibrillary tangles, pathological signatures of AD, cause auditory cortex atrophy. Studies have shown that the volumes of auditory cortex of individuals with AD are significantly smaller than individuals with MCI and healthy controls (Menezes *et al.*, 2025), which affects the central auditory system. Notably, the longitudinal study done in Gates, Beiser, Rees, D'Agostino & Wolf (2002) found that seven out of fifteen individuals with central auditory dysfunction were later diagnosed with probable AD, suggesting that central auditory processing impairment can precede an AD diagnosis. Hence, incorporating routine central auditory evaluations may offer a potential tool for early AD detection.

This evaluation can be achieved through central auditory tests (CATs). CATs are accessible and non-invasive tools crafted for the assessment of the central auditory system. They are auditory tasks that rely on stimuli typically delivered through headphones at a level above peripheral hearing ability. These tests encompass various types, each designed to evaluate different aspects of auditory and cognitive processing. For instance, speech-in-noise tests assess an individual's ability to detect and comprehend speech in noisy environments, such as conversations in a crowded room or background industrial noise (Billings, Olsen, Charney, Madsen & Holmes, 2024). These assessments primarily target selective attention. That is, the ability to concentrate on a specific auditory signal while filtering out distractions. Another category of central auditory tests focuses on reaction time and processing speed, evaluating how quickly an individual responds to auditory stimuli (Tarawneh *et al.*, 2022). These tests provide insight into cognitive reflexes and neural processing efficiency, essential for attentional control and executive function (Niemczak *et al.*, 2021). Notably, the tasks involved in central auditory tests engage cognitive processes more than they assess basic auditory perception, making them valuable for examining central auditory function. This suggests that central auditory assessments could serve as potential indicators of cognitive health and function (Niemczak *et al.*, 2021). Some examples of these tests are listed below:

- **Synthetic sentence identification with ipsilateral competing message test (SSI-ICM)**
  The synthetic sentence identification with ipsilateral competing message (SSI-ICM) is a standardized speech-in-noise test that presents a nonsensical sentence alongside an engaging narrative at the same hearing threshold level (HTL), alternating between ears (Gates *et al.*, 2002). Participants are given a list of sentences and must select the one they hear during each trial. Since each sentence contains multiple distinct keywords, participants can successfully complete the task by identifying just one keyword and matching it to the list. A passing score requires correctly identifying 80% of the sentences. To ensure comprehension, a training phase precedes the test, during which sentences are played at an HTL 10 dB higher than the narrative. The purpose of SSI-ICM is to assess the participant's ability to focus on a specific audio source, distinguish speech, and comprehend spoken content. In Gates *et al.* (2002), researchers used SSI-ICM results alongside a hearing test, where participants identified monosyllabic words in a quiet environment, to identify individuals who performed normally on the hearing test but failed the SSI-ICM. These individuals were diagnosed with central auditory dysfunction, and seven out of fifteen later developed probable Alzheimer's disease.

- **Hearing in noise test (HINT)**
  The HINT is an adaptive speech in noise test that assess the ratio between accurate target speech identification and background noise volume (Nilsson, Soli & Sullivan, 1994). During the test, sentences are administered with a fixed-level background noise (Niemczak *et al.*, 2021). The participant is required to repeat what they hear. The sentence presentation level adjusts dynamically based on performance, if the participant correctly repeats a sentence, the

next one is presented at a higher level, whereas if they fail, the level decreases. HINT engages both ears by varying the noise source direction. For example, the test can be conducted under three conditions, with noise coming from the right, front, and left. Each condition uses a distinct set of sentences (Niemczak *et al.*, 2021). The speech reception threshold (SRT) for each condition is calculated as the average sentence presentation level (excluding the first four), expressed as a signal to noise ratio (SNR). This SNR represents the level at which the participant understands 50% of the words. The combined SRT values from the three conditions form the primary HINT metric for analysis. Findings in Niemczak *et al.* (2021) indicate a strong correlation between HINT performance and cognitive functions such as learning and working memory.

- **Triple digits test (TDT)**

  The TDT is another speech-in-noise assessment in which a recorded sequence of three monosyllabic digits, such as "five, three, two," is played with background noise. In Niemczak *et al.* (2021), the test is conducted in paired presentations with noise. Initially, both the digits and the noise are presented at the same level. However, as the test progresses, the digit presentation level increases by 2 dB SPL when the participant correctly identifies a digit and decreases by the same amount when a digit is missed. The TDT variable is derived from the SNR of the last seven presentations where the positive phase was masked. Findings from Niemczak *et al.* (2021) indicate that TDT is strongly linked to learning and working memory. Furthermore, Niemczak *et al.* (2021) highlights a strong relationship between TDT and the MoCA that evaluates functions such as learning, language, memory, and attention.

- **Dichotic digits test (DDT)**

  The DDT is a widely used assessment of dichotic auditory processing (Tarawneh *et al.*, 2022). It involves simultaneously presenting two monosyllabic digits to each ear at a level of 50 dB above the PTA, and the participant is required to correctly identify the digits heard. According to Gates, Anderson, Feeney, McCurry & Larson (2008), healthy adults typically achieve a 90% success rate on the test. DDT is considered an effective tool for detecting central auditory dysfunction because it is easy to administer, unaffected by mild hearing loss, and produces abnormal results in individuals with AD. In the study conducted in Gates *et al.* (2008), 84% of participants with suspected AD and memory impairments scored below 80% on DDT. This test engages the frontal, temporal, and parietal lobes, as it relies on attention and memory functions (Tarawneh *et al.*, 2022).

- **Dichotic sentence identification test (DSI)**

  The DSI test is another dichotic listening assessment where participants must identify nonsense sentences from a given list (Tarawneh *et al.*, 2022). In DSI, sentences are presented in pairs, with one sentence played in each ear simultaneously. The test offers two modes: in the directed mode, participants identify the sentence in a specified ear, while in the free mode, they must recognize both sentences presented. A standard test consists of five presentations, with an additional five given if the participant does not achieve a perfect score. A score of 80% or higher is considered normal. In Gates *et al.* (2008), 83.8% of participants with suspected AD and memory impairments scored below 80% on DSI. Like the DDT, DSI engages the frontal, temporal, and parietal lobes. Additionally, DSI is linked to the cortical thickness of brain regions associated with language, auditory processing, episodic memory, and working memory. A reduction in cortical thickness in these areas is often considered a marker of AD (Tarawneh *et al.*, 2022).

- **Gap detection test (GAP)**

  The gap detection test (GAP) is an auditory temporal processing assessment that evaluates a participant's ability to perceive changes in sound over a specific duration (Tarawneh *et al.*, 2022). During the test, the participant is required to press a button whenever they detect a gap in the white noise presented. In Niemczak *et al.* (2021), the authors modified the test by implementing an adaptive staircase algorithm, where the gap length increases if the participant fails to detect two consecutive gaps or misses a total of three gaps at a given gap length. This approach aims to determine the participant's gap detection threshold (i.e., shortest gap a subject could reliably detect). However, the GAP test did not show any correlation with the cognitive tests in Niemczak *et al.* (2021). Additionally, auditory temporal processing did not demonstrate significant differences between individuals with MCI and control groups, as reported in Tarawneh *et al.* (2022). As a result, GAP may not be a suitable test for Alzheimer's disease detection.

In summary, CATs, particularly speech in noise tests and dichotic tests, are potential AD detection tools since they assess the state of the central auditory system, which is affected by AD pathology. This impairment is reflected in difficulties with attention and speech perception, linked to the atrophy in the auditory cortex caused by AD. Studies show that results on the CATs suggest that problems in central auditory processing could appear early in the disease and may serve as useful indicators of cognitive decline before more noticeable symptoms develop (Gates *et al.*, 2002).

1.1.3.2 Speech and lexical cues

While CATs have shown to be useful indicators of cognitive decline affecting the auditory cortex, everyday speech and lexical cues have shown promise in reflecting the deterioration in the language and speech areas of the brain due to AD (Ahmed, Haigh, de Jager & Garrard, 2013; Mestach, Hartsuiker & Pistono, 2024), even in its early stages (Hernández-Domínguez, Ratté, Sierra-Martínez & Roche-Bergua, 2018). There are two main regions of the brain involved in language and speech: Wernicke's area and Broca's area (Binder *et al.*, 1997). Wernicke's area is involved in phonological and lexical processing, while Broca's area is responsible for speech production (Flinker *et al.*, 2015). These areas are affected by the brain atrophy caused by AD (Harasty, Halliday, Kril & Code, 1999; Ardila, Bernal & Rosselli, 2016). The picture description task is a non-invasive tool that elicit self-generated speech from which signs of speech and language deterioration can be extracted. Studies have shown that reduced lexical richness, increased pausing, slower speech tempo, and difficulties in finding words are common in individuals with cognitive impairment and AD (Fraser, Meltzer & Rudzicz, 2015; Szatloczki, Hoffmann, Vincze, Kalman & Pakaski, 2015; Toth *et al.*, 2018; Forbes-McKay & Venneri, 2005).

Forbes-McKay, Shanks & Venneri (2013) found that individuals with AD exhibit semantic errors and a decline in the phonological components of speech.

In addition to traditional linguistic analysis, numerous studies in engineering and computer science have been done to detect speech and language impairment due to cognitive decline through natural language processing (NLP) and machine learning methods Nasreen, Hough & Purver; Hernández-Domínguez *et al.* (2018). For example, the research done by Alkenani, Li, Xu & Zhang (2021) used generated speech corpus to develop a machine learning model that combines the predictions of several simpler classifiers to improve accuracy. These classifiers included K-nearest neighbors, support vector machines, logistic regression, decision trees. Each classifier made its own prediction about whether an individual has AD, and then the model combined these predictions to make the final decision. This method of using multiple classifiers and combining them into one is called stacked base classification. The model in Alkenani *et al.* (2021) was tested with written data from AD blog corpus and spoken data from a description of the Cookie Theft Picture. For the written data, the model achieved an accuracy of 97.73%. As for spoken data, the accuracy of the model was equal to 95.35%. Alkenani *et al.* (2021) used lexicosyntactics and n-gram features. Lexicosyntactics focus on the linguistic structure of speech, encompassing grammar and syntax. They include aspects such as sentence length, the total number of words and sentences, the frequency of stopwords (e.g. the, and, a, an), verb usage, and other linguistic features. Individuals with AD often exhibit a decline in language abilities, characterized by repetitive word use and reduced vocabulary. Yancheva & Rudzicz (2016) found that combining lexicosyntactics with acoustic features improves the performance of a random forest classifier in the classification of individuals with AD. N-gram feature spaces represent a sentence by grouping consecutive n words or tokens. For example, in a 2-gram (bigram) representation, the sentence "The student is studying." would be broken into ["the student," "student is," "is studying"]. N-grams can also be formed using character or phoneme tokens. In fact, character-based low-level n-grams are particularly effective for analyzing connected text. Moreover, n-grams tackle both lexical and syntactic information by maintaining the original sequence of words (Wankerl, Nöth & Evert, 2017). Hence, bigrams and trigrams, after removing stopwords, provide valuable features for classification models or statistical analysis. Given the large number of possible features, feature selection techniques are recommended.

The picture description task is not only useful for language and speech studies but it also provides an opportunity to examine eye movements. For instance, Heidarzadeh & Ratté (2023) demonstrated that using the participant's spoken words, they were able to locate the participant's gaze point and build a classifier that identifies AD patients. Hence, adopting a multimodal approach while tackling a unique task offers complementary information and advances the AD detection methods.

1.1.3.3 Eye movements and pupil dilation

Eye movements and pupil dilation have been investigated as potential biomarkers for AD. Among the most studied features are gaze fixation and saccades, which are rapid eye movements from one object to another (Molitor, Ko & Ally, 2015). Studies have shown that individuals with AD have unstable gaze fixation and slower, delayed, and hypometric saccades (Anderson & MacAskill, 2013; Fletcher & Sharpe, 1986; Mosimann *et al.*, 2005). Pupillometry, which is the measurement of pupil diameter over time, provides a complementary AD biomarker by capturing nervous responses. An increase in pupil dilation has been linked to sympathetic nervous system (SNS) activity, which is the "fight-or-flight" response (Alnaes *et al.*, 2014). This response is driven by a neurotransmitter, called noradrenaline, produced primarily by the locus coeruleus (LC), a brainstem nucleus (Stanford, 2001). The LC is among the earliest pathological regions of AD (Heiko Braak, Dietmar R Thal, Estifanos Ghebremedhin & Kelly Del Tredici, 2011; Chen, Chen & Hou, 2022), and since LC activity affects pupil size, pupil responses have been examined as indirect markers of its function (Granholm *et al.*, 2017), and subsequently, a marker for the development of AD (Wu *et al.*, 2025).

1.2 Hearables and their applications in health monitoring

Due to the wide range of physiological and behavioral signals that could serve as early biomarkers for AD, there is a need for devices that can capture them while being non-invasive, accessible, scalable, and accurate. Wearables are portable devices that can be worn on the body, integrated into clothing or as accessories. They are equipped with various sensors capable of measuring different types of data such as biosignals, body movements, in addition to data from the surrounding environment (Xue, 2019). There are several applications for wearables in healthcare due to their non-invasiveness and accessibility. For instance, wristbands and smart watches can monitor heart rate, mental health, and sleep, while chest straps are used to measure heart rate and respiratory activity. Moreover, rings can capture heart rate and peripheral temperature (Dunn, Runge & Snyder, 2018).

A hearable is an intra-aural wearable device that can be equipped with various sensors, including an in-ear microphone (IEM), an outer-ear microphone (OEM) and a speaker. Hearables that strictly rely on acoustic IEMs require an acoustic seal that they build in the ear canal, which attenuates external sounds and creates an occlusion effect. The occlusion effect amplifies low-frequency signals propagated by tissue and bone conduction (Sweetow & Pirzanski, 2003). Due to this amplification, the IEM is able to capture biosignals such as non-verbal audio events. Chabot, Bouserhal, Cardinal & Voix (2021) proposed a real-time machine learning model, bag of audio words (BoAW) that detects and classifies non-verbal audio events such as eye blinks, swallows, and teeth grinding. Various studies were done on extracting heartbeats, that are found below 60 Hz (Martin & Voix, 2018), computing HRs from inside the ear (Benesch, Chabot, Tom, Voix & Bouserhal, 2024; Butkow, Dang, Ferlini, Ma & Mascolo, 2023), and detecting stress (Benesch *et al.*, 2025). Detecting breathing and identifying respiratory phase classification have been also addressed directly with in-ear signals, where Mehrban, Voix & Bouserhal (2024) developed a classifier for breathing phase and path using in-ear signals. Furthermore, Gruters *et al.* (2018) reported that the eardrum moves in synchrony with eye movements, producing oscillations that can be detected by IEM. These oscillations are called eye-movement related eardrum oscillations (EMREOs). Going further, research was done to study the relationship between saccades and EMREOs (Lovich *et al.*, 2023; King *et al.*, 2023). These studies demonstrate the capability of hearables to detect a wide range of physiological and behavioral biosignals that provide information about the sympathetic nervous state of the individual, in addition to being related to various diseases that affect these signals, such as AD.

1.3 Summary

AD is a progressive neurodegenerative disorder that is the most common cause of dementia. Early detection is crucial since it impacts the progression of the disease by providing recommendations for timely clinical care and lifestyle changes that may delay or prevent dementia. Traditional diagnostic methods include expensive and invasive imaging such as MRI and PET scans, in addition to blood tests and cognitive tests. Hence, the need for an accessible, non-invasive, and scalable diagnostic method is on the rise. Research was conducted on various behavioral and physical biosignals that could be potential early biomarkers of AD. Among these, the central auditory processing system, responsible for treating the audio signals that are captured and transmitted to the brain by the peripheral auditory system, is heavily affected by AD. Central auditory tests, such as speech in noise and dichotic tests, assess this system and show that AD patients and people with MCI perform worse on these tests compared to controls. Moreover, biosignals, such as pupil dilation and eye movements, and speech and lexical features, have been studied due to having exhibited differences in AD patients. Table 1.1 summarizes the advantages and disadvantages of the traditional and non-traditional methods for AD. A hearable is an in-ear wearable device that can be equipped with an in-ear microphone that captures all of these biosignals and could also be used to test central auditory function. This makes it a promising tool for monitoring AD as previous studies have extracted and validated heart rate, breathing, and non-verbal audio signals.

Table 1.1 Summary of traditional and emerging non-traditional AD diagnostic methods

Diagnostic methods	**Advantages**	**Disadvantages**
	Traditional	
Imaging and biomarker based	High diagnostic accuracy with high sensitivity and specificity	Expensive, invasive, and limited in accessibility
Cognitive assessment	Highly accessible, low cost, and non-invasive with moderate sensitivity and specificity	Influenced by clinician interpretation and by participant education level
	Non-traditional	
Central auditory tests	Detect central auditory processing deficits; accessible and non-invasive	Must be calibrated to peripheral hearing level; limited clinical standardization for AD detection
Speech and lexical cues	Capture subtle early language and speech changes; natural, non-invasive, and accessible	Affected by language fluency and recording quality; limited clinical standardization for AD detection
Eye movements and pupil dilation	Detect subtle changes in cognitive load and attention; natural and non-invasive	Sensitive to noise (i.e. blinking) and vision corrections; limited clinical validation and standardization

CHAPTER 2

A MULTIMODAL DATABASE FOR DETECTING ALZHEIMER'S DISEASE WITH CENTRAL AUDITORY AND PHYSIOLOGICAL MEASURES: A PILOT STUDY

Miriam Boutros[a,b] , Jenna Stevenson[c,d,e] , Nesrine Rahmouni[c,d,e] , Pedro Rosa-Neto[c,d,e] , Christopher Niemczak[f,g] , Rachel Bouserhal[1,2]

[a] Département de Génie Électrique, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3
[b] The Centre for Interdisciplinary Research in Music Media and Technology, Schulich School of Music, McGill University,
527 Rue Sherbrooke O, Montréal, Québec, Canada H3A 1E3
[c] Translational Neuroimaging Laboratory, The McGill University Research Centre for Studies in Aging, McConnell Brain Imaging Centre,
6825 LaSalle Boulevard, Verdun, Québec, Canada H4H 1R1
[d] Montreal Neurological Institute-Hospital, McGill University,
3801 Rue University, Ville-Marie, Québec, Canada H3A 2B4
[e] Douglas Hospital Research Centre, Centre intégré universitaire de santé et services sociaux de l'Ouest-de-l'Île-de-Montréal, McGill University,
6875 Blvd. LaSalle, Verdun, Québec, Canada H4H 1R3
[f] Dartmouth-Hitchcock Medical Center, Department of Medicine,
1 Medical Center Dr, Lebanon, New Hampshire, United States of America 03766
[g] Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth,
1 Rope Ferry Rd, Hanover, New Hampshire, United States of America 03755

Article submitted to "Hearing Research" in October 2025.

2.1 Abstract

Alzheimer's disease (AD) is the leading cause of dementia, affecting millions of individuals worldwide. Early detection is crucial as it may delay or prevent the progression of cognitive decline. Central auditory tests (CATs), speech and lexical cues, pupillometry and eye movements, and biosignals have shown potential as non-invasive biomarkers of AD. This pilot study presents the Gaze and Auditory Response Database for Evaluating Neurocognitive Impairment and Alzheimer's disease (GARDENIA), a preliminary multimodal dataset collected from 20 participants, including individuals with AD, mild cognitive impairment (MCI), and cognitively unimpaired controls. The database integrates in-ear biosignals, pupillometry and eye tracking, speech and lexical responses, and CATs administered through a hearable. In this report, we

focused on CAT performance. Results indicate that cognitively impaired participants performed worse than controls across all CAT measures, with dichotic tasks showing the strongest predictive value. These findings also demonstrate the feasibility of administering CATs through a hearable while measuring biological signals, highlighting the wearable's multifunctionality and practicality.

2.2 Introduction

AD is the most common form of dementia, responsible for the most dementia cases worldwide (Alzheimer's Disease International, 2022). In Canada, there were over 600,000 dementia cases reported as of 2022, and the number is expected to double by 2030 (Alzheimer Society of Canada, 2023). Similarly, in the United States, there are 7.2 million cases of Alzheimer's disease in individuals aged 65 or older as of 2025 (Alzheimer's Association, 2025a). With the global aging population, the burden of AD is expected to rise substantially, prompting the need for early, scalable, and non-invasive diagnostic tools that can supplement or precede traditional neuropsychological and imaging assessments (Alzheimer's Association, 2025a). Therefore, research has rapidly expanded, leading to the identification of multiple promising measures that may serve as potential markers for early AD detection.

Auditory function has drawn particular attention as a marker for AD (Swords *et al.*, 2018a). Hearing ability naturally declines with age, but the brain's ability to process complex auditory sounds may also provide insight into neurocognitive changes associated with AD. While peripheral auditory function (i.e. raising your hand to a pure tone) has been reported to have mixed correlations with cognitive ability (Lin *et al.*, 2011; Thomson, Auduong, Miller & Gurgel, 2017). More complex auditory tasks, termed central auditory processing, such as speech perception in background noise, have shown to reflect diverse neurocognitive domains (Gates *et al.*, 2008). The central auditory system (i.e. beyond the cochlea) is primarily responsible of the execution of these complex tasks by engaging both cortical and subcortical pathways (Felix, Gourévitch & Portfors, 2018). Particularly, studies have shown that central auditory dysfunction does not only correlate with AD, but it may also precede the onset of clinical dementia (Gates *et al.*, 2002, 2011). This may be explained by the development of amyloid plaques and neurofibrillary tangles in the temporal lobe (Therriault *et al.*, 2022), which results in difficulty understanding speech in background noise, reduced attention, and decline in memory, speech, and language functions (Holmes, 2014).

CATs are designed to assess the auditory system beyond the cochlea and include a variety of measures. HINT and TDT are two popular central auditory tests that assess speech perception in noise and correlate with learning and working memory function (Niemczak *et al.*, 2021). Dichotic tests, such as DDT and DSI, involve identifying different messages presented simulataneously in both ears and have shown to correlate with attention and memory function (Lipschutz, Kolinsky, Damhaut, Wikler & Goldman, 2002). Literature has shown that AD patients and individuals with cognitive impairment perform worse on tasks requiring central auditory processing than age-matched controls (Tarawneh *et al.*, 2022).

Daily, ongoing speech and lexical components have also emerged as a source of information for AD detection. Language impairments, such as simple lexical complexity and slower speech tempo are some of the earliest signs of AD (König *et al.*, 2015; Petti, Baker & Korhonen, 2020; Ortiz, De Lira, Minett & Bertolucci, 2024). Tasks that elicit self-generated speech, such as picture description, provide valuable tools for analyzing linguistic features including lexical diversity, semantic content, fluency, brevity, and syntactic complexity (Mueller, Hermann, Mecollari & Turkstra, 2018). Studies have shown that reduced lexical richness, increased pausing, slower speech tempo, and difficulties in finding words are more common in individuals with cognitive impairment and AD (Fraser *et al.*, 2015; Szatloczki *et al.*, 2015; Toth *et al.*, 2018). Previous research has used NLP techniques to extract, quantify, and analyze these speech and language features to detect AD (Alkenani *et al.*, 2021).

With advances in wearable technology, research on in-ear devices as tools to develop AD biomarkers is increasing (Seol & Moon, 2022; Lin *et al.*, 2011; Musaeus *et al.*, 2022). These devices are characterized by their non-invasive nature and the ease of integration into the wearer's daily routines and environments. Particularly, in-ear devices called hearables that are equipped with IEMs, capture biosignals from the ear such as heartbeats, breathing, eye blinking, and swallowing (Röddiger *et al.*, 2022) using only the IEM. This is possible due to the occlusion effect that occurs when the ear canal is sealed at its opening and amplifies low-frequency signals propagated by tissue and bone conduction (Wang, Lu, Sang, Cai & Zheng, 2022). Previous studies employed IEMs to detect and classify non-verbal biosignals (Chabot *et al.*, 2021), stress (Benesch *et al.*, 2025), breathing phases (Mehrban *et al.*, 2024; Martin & Voix, 2018), and eye movements (Gruters *et al.*, 2018; Lovich *et al.*, 2023). These signals offer a passive and continuous way to assess autonomic nervous system activity (Joshua A. Waxenbaum, Vamsi Reddy & Matthew A. Varacallo.), which is increasingly recognized as being disrupted in neurodegenerative conditions (Jandackova, Scholes, Britton & Steptoe, 2024; Liu, Elliott, Knowles & Howard, 2022). However, despite the increasing research in in-ear signal processing, studies directly assessing cognitive decline remain limited, especially with populations with diagnosed AD.

Finally, beyond speech and language, other non-invasive modalities have also been investigated to capture neurocognitive changes. Eye tracking and pupillometry offer non-invasive windows into attention, arousal, and working memory. Eye movement patterns, including fixation and saccades, have also been found to differ in AD patients. AD patients have less stable fixation and slower, more delayed, hypometric saccades (Anderson & MacAskill, 2013; Fletcher & Sharpe, 1986; Mosimann *et al.*, 2005). Additionally, pupil dilation during cognitive tasks have been associated with locus coeruleus (Alnaes *et al.*, 2014). Locus coeruleus is a small nucleus in the brain stem responsible for the production of 45% of noradrenergic neurons that release noradrenaline (Stanford, 2001), the neurotransmitter involved in the “fight-or-flight response” and attention (Sheppard *et al.*, 2024). The noradrenaline released by locus coeruleus activity increases the pupil diameter and suppresses the parasympathetic system (Alnaes *et al.*, 2014). AD causes degeneration of locus coeruleus neurons and reduction in noradrenaline (Chen *et al.*, 2022). Studies have shown that this nucleus is one of the earliest affected regions in AD pathology (Heiko Braak *et al.*, 2011) and may serve as a biomarker for early detection of AD (Wu *et al.*, 2025; Iannitelli *et al.*, 2023; Al Haddad *et al.*, 2023; Sibahi *et al.*, 2023). Therefore, since pupil responses are linked to locus coeruleus activity, studying pupil dilation might offer a non-invasive tool to gauge the state of locus coeruleus and highlight the presence of AD (Granholm *et al.*, 2017).

The first goal of this study is to present the GARDENIA, a multimodal database collected from individuals diagnosed with AD, MCI, and cognitively unimpaired controls. GARDENIA focuses on four major components: in-ear biosignals, pupillometry and eye movements, speech and lexical responses, and CATs. Each component of the dataset was selected based on emerging evidence linking it to AD and health monitoring. The purpose of this database is to contribute to the growing field of continuous health monitoring and early disease detection through accessible and non-invasive technologies. While research has been done on the four potential biomarkers individually, there is lack of data tackling these components together. Hence, GARDENIA was created to provide this type of data by collecting speech and biosignals simultaneously while doing CATs.

The second goal of this study is to analyze one component of the database to assess its utility. The CATs were selected for this purpose because literature has proven, recurrently, that people with AD perform worse on the tests. Thus, this goal focuses on analyzing the results on the CATs to determine whether the test results are consisted with findings reported in literature, as in whether they can differentiate cognitively impaired (CI) from the cognitively unimpaired (CU) participants. The analysis also tackles the identification of the type of tests that have the highest AD predictive value. Finally, CATs results provide validation for the methods used in the data collection, which provides a preliminary assessment of the feasibility of administering these tests through a hearable. This aspect of the analysis offers insight into the potential for real-world deployment using a non-invasive, accessible, and multi-purpose platform.

In this report, the data collection procedure is presented in section 2.3 along with a description of GARDENIA in section 2.4. Analysis and results of the CATs are presented in section 2.5 and section 2.6, respectively, and discussed in section 2.7.

2.3 Methodology

The project was approved by the *Comité d'éthique de la recherche*, the ethics review board, of the École de technologie supérieure (25080) and the *CIUSSS de l'Ouest-de-l'Île-de-Montréal* (IUSMD-21-61). All participants read and signed a consent form that was approved by the ethics committees.

2.3.1 Participants

Participants were recruited through the Translational Biomarkers of Aging and Dementia (TRIAD) cohort. Each participant had the diagnosis of Alzheimer's disease confirmed by amyloid and tau PET scans, a complete battery of behavioral neurocognitive tests, including the CDR, the MoCA, and the MMSE (Nasreddine *et al.*, 2005; Berg, 1984; Folstein *et al.*, 1983). Of 20 individuals (65-80 y.o.) recruited, eight met criteria for cognitive impairment due to AD based on cognitive measures (MMSE < 24 and CDR = 1) and positivity on brain amyloid and tau PET scans (CI; N = 8). Twelve CU individuals performed on the normal range of cognitive assessments, scored 0 on the CDR, and had a negative PET scan for amyloid or tau (CU). All participants completed a hearing health questionnaire either in person or by phone during the recruitment process. The data collection was conducted at the McGill University Research Centre for Studies in Aging (MCSA). Three peripheral auditory examinations were performed followed by four central auditory tests and a picture description task. Table 2.1 outlines the protocol followed during the data collection.

2.3.2 Instruments

To capture in-ear signals, a custom hearable was used, as presented in Figure 2.1. It contains an IEM, an OEM, and in-ear speaker in each ear. Foam Comply™ Original Ear Tips (Hearing Components, Inc., Oakdale, Minnesota, United States) were used for passive noise attenuation. Component specifications and more technical information could be found in Bouserhal, Bernier & Voix (2019).The audio signals were recorded at a sampling rate of 48000 Hz.

Table 2.1 Data collection protocol showing the test battery and the estimated time in minutes for each test

Test Battery Components	**Estimated Time (minutes)**
Informed consent briefing	10
Otoscopy Tympanometry Audiometry	5 (total for both ears) 5 (total for both ears) 10 (total for both ears)
Wearing the sensors Fit check	5 10
Recording starts	
Finding MCL Hearing in noise test Triple digits task Dichotic digits test Dichotic digits identification test Picture description task	5 15 15 10 5 5
Recording ends	
Total:	**100 minutes**

Figure 2.1 Hearable used in data collection

To track eye movements and pupil dilation, Tobii Pro Glasses 3 (tobii, Stockholm, Sweden) were used set at a sampling rate of 100 HZ. Detachable lenses were clipped to the glasses for participants with corrected vision. To acquire heart rate, an earlobe clip PPG from Shimmer GSR+ (Shimmer, Dublin, Ireland) was used. PPG measures the blood volume from which the heart rate can be computed. Moreover, skin conductance was measured using a finger GSR sensor from the same Shimmer GSR+ toolkit. The PPG and the GSR data are available for only the first four participants due to malfunction in the sensor.

The speakers of the hearable were connected to a Rolls HA43 Pro amplifier (Rolls Corporation, Salt Lake City, Utah, United States) which was connected to a Roland OCTA-CAPTURE (Roland Corporation, Hamamatsu, Shizuoka Prefecture, Japan) soundcard during the data collection. The microphones were connected to the soundcard only for the fit test that was done at the start of the session to determine whether the earpiece was correctly inserted and whether the four microphones were functioning properly. The speaker Genelec 8030C (Genelec Inc, Natick, Massachusetts, United States) was used to play a loud 80 dBA pink noise during the fit test. Then, during the following tests, the IEMs and OEMs were connected to a ZOOM H6 (Zoom North America, Hauppauge, New York, United States) hand recorder that was placed in front of the participants. The ZOOM H6 two-channel microphone (90° angle) was placed 15 cm in front of the mouth and used as a reference microphone. The Tobii Pro glasses, the Shimmer GSR+, and the soundcard were connected to a laptop that was using lab streaming layer (Chadwick Boulay & Tristan Stenner, 2025; Kothe *et al.*, 2025) applications to synchronize the two sensors with markers sent from a MATLAB 2023a (MathWorks, Natick, Massachusetts, United States) graphical user interface pertaining to the central auditory tests.

2.3.3 Peripheral auditory tests

At the start of the data collection three peripheral auditory examinations were conducted:

2.3.3.1 Otoscopy

Otoscopy was performed on both ears to assess the presence of cerumen within the ear canal. Participants with occluded ear canals (i.e. complete blockage) were excluded to ensure that the IEM could properly capture the biosignals since cerumen forms a physical barrier in the ear canal.

2.3.3.2 Tympanometry

Tympanometry was conducted using the Maico EasyTymp (MAICO Diagnostics US, Minnesota, United States) tympanometer to evaluate the movement and the compliance of the tympanic membrane, as well as assess cerumen occlusion if noted during otoscopy. Therefore, participants with abnormal tympanometry were excluded, but no participant displayed this configuration. All tympanometry measures were reviewed by an audiologist.

2.3.3.3 Audiometry

Audiometry was performed using the SHOEBOX PureTest (SHOEBOX, Ottawa, Ontario, Canada). The purpose of the test is to find lowest hearing volume a participant could reliably detect (i.e. threshold) at frequencies 500, 1000, 2000, 4000, and 8000 Hz. Participants were excluded if they had a PTA greater than 40 dB HL in either ear, computed over the 500, 1000, 2000, and 4000 Hz frequencies.

2.3.4 Picture description task

In this task, the participants were asked to describe the picture on the screen in front of them. Two pictures were presented with the first one being “The Boston Cookie Theft Picture" (Harold Goodglass, Edith Kaplan & Barbara Barresi) and the second one being “The Traffic Chaos Picture" taken from Forbes-McKay & Venneri (2005); Forbes-McKay *et al.* (2013).

2.3.5 Central auditory tests

2.3.5.1 Presentation level through the hearable

Central auditory tests were conducted through the hearable instead of a traditional headset used in previous literature. An artificial head was used to measure the level of presentation of stimuli in dBA through the hearable's in-ear speakers in the ear canal. The amplifier was used to avoid clipping for high levels (e.g. 90 dBA). Using the OCTA-CAPTURE at a fixed gain of 11 dBA and the laptop at 100% volume, a calibration factor indicating the difference between the required level of presentation and the actual level inside the artificial head was computed for every stimulus at 65 dBA. Hence, six calibration factors were computed for speech shaped noise, babble noise, HINT sentences, TDT sets, DDT sets, and DSI sentences. Equation (1) shows how the calibrated most comfortable level (MCL) was computed based on the calibration factor measured using the artificial head and the most comfortable level (MCL) that was personalized to every participant.

$$calibratedMCL = calirationfactor + MCL - 65 \tag{2.1}$$

Equation (2) shows how the signal was amplified based on the calibrated MCL and the multiplier, which indicates the number of decibels that needed to be added or subtracted from the presentation levels in the speech in noise tests.

$$\text{speech}_{\text{amplified}} = \text{speech} * 10^{(\text{multiplier} + \text{calibrated MCL})/20} \tag{2.2}$$

The plots in Figure A I.1 in section I show the linear relationship between the compensation (i.e. calibration factor for each stimulus and the presentation level at the artificial head. This linear relationship means that for every one unit increase in the calibration factor, the level of presentation increases one dBA at the artificial head. Hence, this relationship ensures that the

CATs are properly administered through the hearable by successfully controlling the presentation level of the stimuli.

2.3.5.2 Most comfortable level

To account for the peripheral hearing loss due to aging and account hearing ability on central auditory tests, a MCL was found. This level was obtained by adding the highest PTA from the audiometry results to 60 dBA. Sentences were then presented binaurally at this level, and participants were asked to repeat what they heard. If the sentence was repeated correctly, the presentation level was reduced by 10 dBA. This process continued until the participant could not repeat 50% of the words in a sentence. At that point, 30 dBA was added to the level found. Participants were then asked whether this level felt comfortable as in comparable to typical listening levels for television or radio, and whether they would prefer to adjust it. If the participant wanted the level louder or softer it was adjusted by 5 dBA increments. Once a participant confirmed the level as comfortable, it was established as their MCL.

2.3.5.3 Central auditory tests

Hearing in noise test (HINT): Four lists of 20 sentences each were presented in both ears with a background noise. The background noise for lists 1 and 3 was speech-shaped noise (i.e. an energetic masker), while for lists 2 and 4, was an eight-talker French babble noise (i.e. informational masker). The presentation level of the noise was fixed at the MCL. The first sentence was presented at 6 dBA SNR to familiarize the participants with the test. That is, the sentences were presented 6 dBA louder than the background noise. The participants were asked to repeat what they hear. The presentation level of the sentences was increased by 2 dBA when the participants did not successfully repeat all the words. On the other hand, it was decreased by 2 dBA when the sentence was correctly repeated. The mean SNR for each of the speech shaped noise and babble noise, excluding the first six presentations was used as the primary outcome variable.

Triple digits test (TDT): Two lists of 30 sets of digit triples were presented in both ears with background noise. Similar to HINT, the first list had speech-shaped noise while the second list had babble noise. The noise was also fixed at the MCL. The first set of digits was presented at 6 dBA SNR and the participants were asked to repeat the digits they heard. The presentation level of the sentences was increased by 2 dBA when the participants were not able to repeat all the three digits, otherwise, it was decreased by the same amount. Similarly to HINT, the mean SNR for each of the speech shaped noise and babble noise, excluding the first six presentations was used as the primary outcome variable.

Dichotic digits test (DDT): 25 presentations of two-digit pairs were simultaneously played in each ear (i.e. total of four distinct digits in each presentation) were presented at the MCL. The participants were then asked to repeat the four digits in no distinct order. The number of correct full presentation (i.e. successful identification of the stimuli in both ears) was computed as the primary outcome variable.

Dichotic sentence identification test (DSI): Two lists of ten presentations each were presented at the MCL, while a list of ten sentences was shown on a computer screen in front of the participants. Two different sentences from the list on the screen were simultaneously played in each ear. The sentences, taken from Tuwaig *et al.* (2017), were nonsense French sentences that do not follow any syntax or linguistic rules. The participants were asked to identify the two sentences that were played. The number of correct identifications of presentations was calculated as the primary outcome variable.

2.4 Database

The GARDENIA database contains the following data from 20 participants:

- Basic demographic data including age, gender, years of education.
- Answers to a health questionnaire about their hearing health and habits in French.
- Cognitive diagnosis based on the results of various cognitive assessment tests that was assembled by Douglas Mental Health Institution. The data entails the results on the Montreal

Cognitive Assessment (MoCA) Nasreddine *et al.* (2005) test that vary between zero and 30, with 26 and above indicating normal cognitive state. The scores on the Mini Mental State Examination (MMSE) (Folstein *et al.*, 1983) test that range from zero to 30, with 24 and above signaling normal cognitive state. Finally, the ratings on the clinical dementia rating (CDR) (Berg, 1984) scale that range from zero to three, with three indicating severe dementia.

- The results of the otoscopy exam that determine whether the ears were clean, blocked, or have earwax.
- The tympanometry plots.
- The audiometry results for the 500, 1000, 2000, 4000, and 8000 Hz frequencies.
- The fit check recording from which the coherence and the attenuation of the hearable can be extracted.
- The most comfortable level.
- The results of HINT and TDT as presentation levels in dBA and whether there was an increase or decrease of 2 dBA between each presentation.
- The results of DDT and DSI as binary values for each individual, with 1 indicating a correct identification of the stimulus.
- The audio recording of the picture description task and the four CAT in French. This audio data includes the simultaneous recording from 5 different microphones: the reference microphone placed in front of the participants, the left and right channel of the OEM, as well as the left and right channel of the IEM.
- Eye-related data that includes the pupil diameter for each eye, eye movements shown by the gaze data, and gyroscope data to track the position of the head.
- PPG signal taken from an earlobe clip and GSR data from finger sensors for only four participants.
- Given the hardware setting, markers were used to synchronize the audio data with the eye and PPG data. These markers are stored along with the eye and PPG data in one .xdf file containing the appropriate timestamps. A map that explains the meaning of each marker is also provided.

2.5 Analysis

In this exploratory study, statistical analysis was focused on the results of the CATs and their relationship with cognitive impairment. Z-scores were used to compute three composites: speech-in-noise tests, which comprise HINT and TDT, dichotic tests, which include DDT and DSI, and a global CAT that encompasses all four tests.

Z-scores were calculated using the CU group as the reference. For each participant, this involved subtracting the mean of the CU group from their score and then dividing by the standard deviation of the CU group. For the DDT and the DSI z-scores, a positive z-score value indicates that the participant performed better than the average CU participant, and a negative score indicates that the participant performed worse than the average CU participant. The opposite is true for the HINT and TDT z-scores where a positive z-score means that the participant performed worse than the average CU participant having a higher SNR instead of a lower one. Because the z-scores for dichotic tests and speech in noise tests have opposite interpretation of the same direction, the z-scores of HINT and TDT were multiplied by -1. This sign inversion of the speech in noise tests ensures that higher z-scores referred to better performances for all auditory tests.

Analysis of variance (ANOVA) was performed using raw scores to evaluate the difference in performance between the CI and CU groups. Because the CI and CU groups were not matched by age, the CAT results were adjusted to account for the effect of age. The adjustment consisted of fitting a linear model to the scores of each of the six tests using only the data from the CU group. The predicted scores of the two groups were then computed using the fit line of the CU group. The difference between the actual score and the predicted score comprised the age-adjusted scores. Therefore, assuming that there is a linear relationship between the scores and age as plotted in Figure 2.3, we were able to compare CI participants to CU participants without the effect of age.

Three logistic regressions were also performed to evaluate the effect of the three age-adjusted CATs composites on predicting whether the participant is impaired, while also controlling for

PTA. Due to the small sample size, logistic regression was performed using only two predictors at a time (CAT score and PTA). Age-adjusted scores were used instead of raw scores since the two groups are almost perfectly split by age (see demographics in Table 2.2). Despite administering the tests using the most comfortable level belonging to each of the participants, PTA was added to the model to further eliminate any biases pertaining to peripheral hearing ability.

We set the significance threshold at $p < 0.05$, fully recognizing that this approach does not correct for potential inflation of type-I error due to multiple comparisons. Given the exploratory nature of this analysis and the relatively small sample size, this risk was accepted to identify potentially meaningful patterns that warrant further investigation in larger confirmatory studies with this evolving database.

2.6 Results

Table 2.2 presents the demographics of the participants. The impaired group had a mean age of 68 years, with three females and five males. In comparison, the unimpaired group had a lower mean age of 59 years, including seven females and five males. Peripheral hearing ability was measured using the PTA which was calculated using the means over frequencies of 500, 1000, and 2000 Hz. The t-tests show that only age significantly differed between the two groups unlike gender, years of education, and PTA.

Figure 2.2 shows a boxplot illustrating the performance of CI and CU groups on the tests. The detailed scores can be found in Table A I.1 in Appendix A. The two groups had different means on all tests with the CI group performing worse. This is shown in the results of the ANOVA tests (Table 2.3) that revealed a statistically significant effect of cognitive diagnosis across all individual central auditory tests and composite measures. Specifically, the CI group performed significantly worse than the CU group on the HINT in speech-shaped noise ($F = 6.12$, $p = 0.025$), HINT in babble noise ($F = 6.17$, $p = 0.024$), TDT in speech-shaped noise ($F = 10.90$, $p = 0.004$), and TDT in babble noise ($F = 8.02$, $p = 0.012$). Strong effects were also observed for dichotic

Table 2.2 Distribution of demographic features in the CI and CU groups, in addition to the results on the Montreal Cognitive Assessment (MoCA) test, the Clinical Dementia Rating (CDR) scores, and the results of the Mini Mental State Exam (MMSE). P-values refer to the results of t-tests for continuous features, except for the percentage of females, which was assessed using a chi-square test

	CI	**CU**	**P-values**
Number (% of the total)	8 (40%)	12 (60%)	-
Age (SD) in years	67.87 (5.222)	58.83 (4.218)	**<0.001**
Female (% of group)	3 (38%)	7 (58%)	0.361
Education (SD) in years	16.37 (4.438)	15 (3.593)	0.455
Pure tone average (SD) in dB HL	14.25 (4.318)	11.5 (3.618)	0.140
MoCA	21.25 (5.092)	28.17 (1.467)	-
CDR	0.688 (0.259)	0	-
MMSE	26.00 (3.207)	29.67 (0.651)	-

tests, including the DDT ($F = 31.08$, $p < 0.001$) and DSI ($F = 10.34$, $p = 0.005$). Composite scores further supported these findings, with significant effects observed for the speech-in-noise tests ($F = 12.88$, $p = 0.002$), the dichotic tests ($F = 22.11$, $p < 0.001$), and the global CAT score ($F = 18.19$, $p < 0.001$). In contrast, neither Age nor PTA showed significant main effects on CAT performance (all $p > 0.05$), suggesting that the group differences observed were primarily driven by cognitive diagnosis rather than peripheral hearing ability or age.

As for the adjusted scores, the CI group had a mean SNR of -1.03 dBA(SD = 1.42) for HINT SSNoise, -1.80 dBA (SD = 3.03) for HINT babble noise, -1.92 (SD = 2.20) for TDT SSNoise, and -2.33 (SD = 3.04) for TDT babble noise. For the dichotic tests, the CI group had a mean -11.29 (SD = 6.05) of correctly identified presentations for DDT and -6.31 (SD = 6.54) for DSI. Since the adjustment was done using the regression line for CU, the adjusted scores for the CU group are close to zero in all tests. Figure 2.3 shows the predicted scores for the two groups using the linear regression model that accounts for age. For both noise conditions in HINT and TDT, the regression line had a positive slope indicating that the hearing threshold increases with age for the CU group. The scatter plots show that the points above the linear line belong to poorer performances, while the points below the line refer to better performances. For the two dichotic tests DDT and DSI, the linear regression slope was negative, indicating that the points

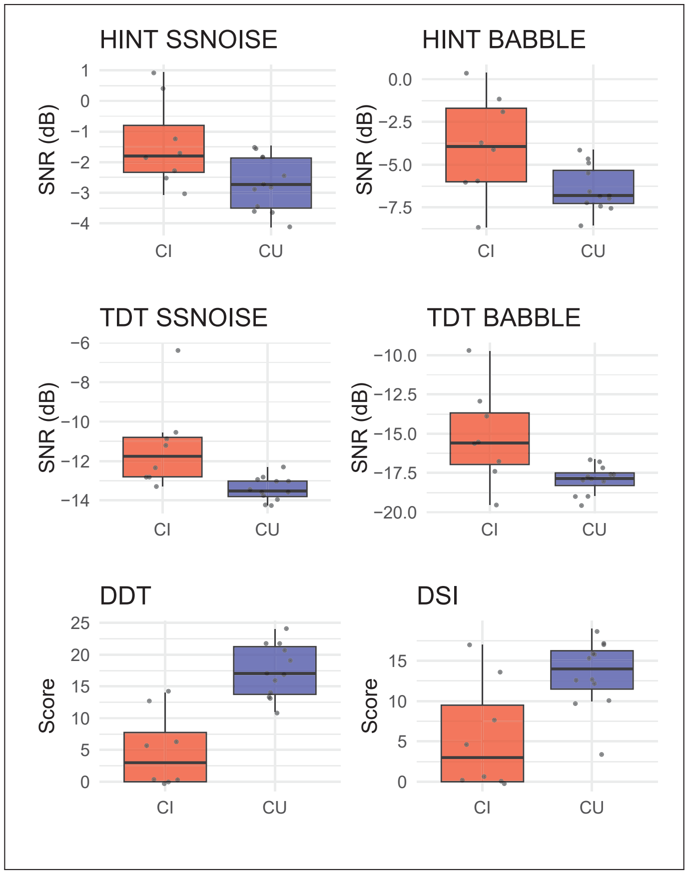


Figure 2.2 Boxplot showing the performance of the CI and the CU groups on the CATs.Higher values indicate poorer performance on the HINT and TDT, while lower values indicate poorer performance on the DDT and DSI

lying below the line belong to the worse performances, while the points above the line signify better performances. In all four tests, the majority of the CI group performed worse than the predicted CU group with similar age.

The results of the three logistic regression models are shown in Table 2.4. The first model M1, which includes the composite speech in noise that combines HINT and TDT adjusted scores, has a moderate fit (AIC = 23.14, R2 = 0.36). For every unit increase in adjusted speech in noise scores, the log odds of being impaired decrease by 1.37 with an odds ratio of 0.25 (p = 0.069), controlling for PTA. However, because of the p-value and the confidence interval [0.06,

Table 2.3 Analysis of variance of the effects of diagnosis, age, and PTA on the performance on the central auditory tests and the three composites

	Diagnosis	Age	PTA
HINT SSNoise	F = 6.12 **p = 0.025***	F = 0.08 p = 0.770	F = 1.63 p = 0.277
HINT Babble	F = 6.17 **p = 0.024***	F = 0.02 p = 0.868	F = 0.30 p = 0.546
TDT SSNoise	F = 10.90 **p = 0.004****	F = 0.28 p = 0.599	F = 2.71 p = 0.119
TDT Babble	F = 8.02 **p = 0.012****	F = 0.02 p = 0.878	F = 0.14 p = 0.708
DDT	F = 31.08 **p < 0.001*****	F = 0.007 p = 0.933	F = 2.18 p = 0.159
DSI	F = 10.34 **p = 0.005****	F = 1.29 p = 0.273	F = 1.47 p = 0.242
Speech in noise tests (HINT + TDT)	F = 12.83 **p = 0.002****	F = 0.06 p = 0.800	F = 0.73 p = 0.404
Dichotic tests (DDT + DSI)	F = 22.11 **p < 0.001*****	F = 0.44 p = 0.512	F = 0.01 p = 0.915
CAT (HINT + TDT + DDT + DSI)	F = 18.19 **p < 0.001*****	F = 0.20 p = 0.656	F = 0.22 p = 0.641

Table 2.4 Results of the logistic regression that assess the ability of the three composites to predict the diagnosis of the participants

		Estimate	Odds Ratio	p-value	95% CI lower	95% CI upper	AIC	McFadden's R^2
M1	Intercept	-3.16	0.042	0.16	0.0005	3.60	23.14	0.36
	Adjusted speech in noise (HINT + TDT)	-1.37	0.25	0.069	0.06	1.11		
	PTA	0.14	1.15	0.36	0.80	1.59		
M2	Intercept	-3.63	0.026	0.12	0.00006	1.36	19.70	0.49
	Adjusted dichotic (DDT + DSI)	-1.62	0.19	**0.031***	0.02	0.59		
	PTA	0.14	1.15	0.38	0.9	1.74		
M3	Intercept	-3.42	0.033	0.14	0.00009	1.72	20.87	0.44
	Adjusted CAT (HINT + TDT + DDT + DSI)	-1.63	0.19	**0.045***	0.02	0.63		
	PTA	0.14	1.15	0.38	0.90	1.72		

1.11], the adjusted speech in noise scores composite was not a significant negative predictor of cognitive impairment.

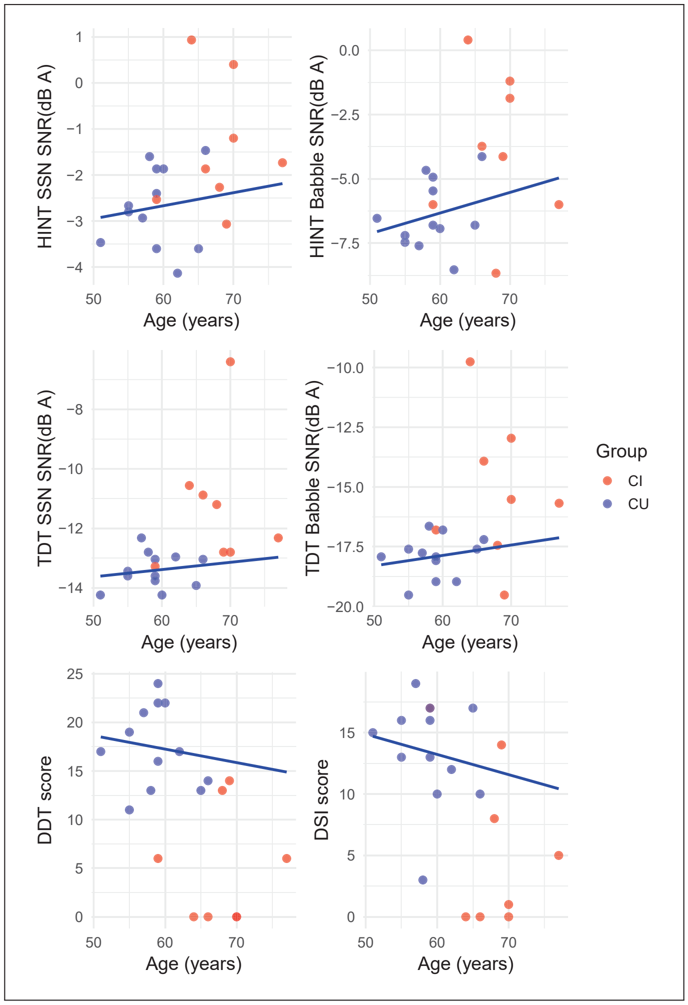


Figure 2.3 The predicted scores on the central auditory tests of the CI (red) and CU (blue) groups based on the linear regression using CU scores along with age. The blue line for each test represents the linear regression line fitting the CU data points. For speech-in noise tests, better performance is represented by lower SNR. As for dichotic tests, greater scores indicate better performance

The second model M2, which includes the composite dichotic DDT and DSI adjusted scores, had the best fit (AIC = 19.70, R2 = 0.49). For every unit increase in adjusted dichotic scores, the log odds of being impaired decrease by 1.62 with an odds ratio of 0.19 (p = 0.031), controlling for PTA. Hence, the adjusted dichotic scores composite was a significant negative predictor of cognitive impairment.

The third logistic regression model M3 combined all four CATs into one composite adjusted CAT score and demonstrated strong fit (AIC = 20.87, R2 = 0.44). For every unit increase in adjusted CAT scores, the log odds of being impaired decrease by 1.633 with an odds ratio of 0.19 ($p = 0.045$), controlling for PTA. Therefore, adjusted CAT scores composite was a nominally significant negative predictor of cognitive impairment.

In all three models, PTA remained a non-significant positive predictor of cognitive impairment ($p>0.05$).

2.7 Discussion

This study corroborates previous literature suggesting that cognitively impaired patients have weaker central auditory processing functions and subsequently perform worse on central auditory tests (Tarawneh *et al.*, 2022). This was shown in the ANOVA results where participants with cognitive impairment consistently demonstrated poorer performance compared to cognitively unimpaired individuals across all four central auditory tests. Furthermore, age and peripheral hearing ability indicated by the PTA did not present any significant effect on the difference in performance.

The results of the logistic regression indicated that central auditory tests were significant predictors of cognitive impairment due to Alzheimer's disease, which was not the case for PTA. However, the significance of the effect of the tests was mostly carried by dichotic tests which suggests that the cortical regions related to dichotic auditory performance are heavily affected. It is worth noting that many CI participants were able to correctly identify the stimuli played in only one ear, notably the right one, which is also reported in literature (Tarawneh *et al.*, 2022; Thamizhmani, Ganapathy, Palaniswamy, Pitchaimuthu & Adhikari MR, 2025a). According to Bouma & Gootjes (2011), dichotic tests in free recall mode engage bottom-up processing of verbal information, which depends significantly on the brain's hemispheric asymmetry in language and auditory processing often leading to better performance in the right ear. Moreover, declines in attention and executive functioning, due to frontal lobe pathology changes in AD patients, lead to more difficulty processing stimuli presented to the left ear compared to the right (Gootjes *et al.*, 2006). In contrast, despite being significantly different between the two groups, HINT and TDT were not significant predictors of cognitive impairment in AD patients. Additionally, the results show that the two groups perform better with babble noise than the speech-shaped noise. This observation contradicts previous study stating that babble noise affects cognitive processing compared to speech-shaped noise affecting more peripheral structures (Vander Werff, Niemczak & Morse, 2021). However, this study used eight-talker babble which could limit the information masking effect (i.e. lower number of talkers equate to more informational masking). This may be a reason for this contradiction since the babble noise comprises eight talkers. A larger sample number is needed to assess the significance and the nature of this finding.

In addition to identifying CI participants, the results of this study also highlight the feasibility of administering CATs through an in-ear wearable device. The implementation of hearing in noise and dichotic tests was successful given that it was able to reproduce the results in literature which indicates that there was a significant difference in performance between CI and CU groups, where the CI group performed worse than the CU group. This contribution is particularly advantageous as it enables these tests to be conducted in the field, outside of specialized clinical environments, and without requiring trained personnel for administration. Moreover, since hearing aids are already commonly used among the aging population, integrating such testing capabilities using hearable technology into everyday life is a natural and practical step. With ongoing research into capturing and analyzing biosignals from hearables, these devices may in the future serve as potential multifunctional tools for auditory assessment, health monitoring, hearing aid, and noise cancelling. Apart from technological innovations, one of the strengths of this pilot is that it relies on a well-characterized population, with detailed clinical characterization and identification of AD neuropathology using second-generation PET tracers for amyloid and tau. Note that nearly 30% of clinical diagnoses of AD do not have amyloid and tau in the brain (Doody *et al.*, 2014).

There were several limitations that affected the study. The small sample size limited generalizability and the ability to perform more powerful statistical tests on the results. Hence, future research should include a bigger cohort where the level of cognitive impairment can be studied since cognitive impairment varies on a spectrum including mild cognitive impairment, AD, and other cognitive diseases, such as Parkinson's disease, across all of their respective stages. Moreover, the CI and the CU groups were not age-matched, which affected the analysis. Although the scores were age adjusted, the analysis is still not generalizable since the adjustment was done assuming linearity between performance and age. Finally, there is a risk of multiple comparisons due to the number of statistical significance tests performed.

2.8 Conclusions

The results of the study support the use of central auditory tests, particularly dichotic auditory tests, as a non-invasive and accessible tool for detecting cognitive impairment in individuals with Alzheimer's disease. Moreover, during this study, these tests were administered through a hearable, which indicates that the tests can be done using an in-ear wearable device while maintaining their ability to identify cognitively impaired individuals. This aspect of the study contributes to the applications of in-ear wearable devices to aid in auditory testing and potentially measuring daily automatic speech recognition, hearing aid, and noise canceling. Finally, the study contributes to the field of cognitive impairment and Alzheimer's disease detection by building a database that contains multimodal data from individuals with Alzheimer's disease or mild cognitive impairment. While this report tackles one aspect of the database pertaining to the relationship between the central auditory function and Alzheimer's disease, future work can be done by validating the biosignals captured by the in-ear microphone, analyzing eye movements and pupillometry data, and examining the speech produced in the picture description task.

2.9 Funding

This study was funded by ÉTS Marcelle Gauvreau Engineering Research Chair in multimodal health monitoring and early disease detection.

The TRIAD cohort study is supported by the Weston Brain Institute, Canadian Institutes of Health Research (CIHR) [MOP-11-51-31; RFN 152985, 159815, 162303], Canadian Consortium of Neurodegeneration and Aging (CCNA; MOP-11-51-31-team 1), Brain Canada Foundation (CFI Project 34874; 33397), and the Fonds de Recherche du Québec – Santé (FRQS; 2020-VICO-279314; 2024 VICO-356138; https://doi.org/10.69777/324345 ; https://doi.org/10.69777/356138 and https://doi.org/10.69777/312994).

2.10 Acknowledgement

The authors would like to acknowledge Siyana Milenova Milanova for synchronizing the sensors using Lab Streaming Layer. The authors would like to acknowledge the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) for lending Tobii Pro Glasses 3 and the speaker and Dr. Suresh Krishna for lending the lenses for Tobii glasses. Finally, the authors would like to acknowledge Dartmouth College for lending the otoscope and tympanometer, and Valentin Pintat for troubleshooting hardware issues.

CHAPTER 3

PRELIMINARY RESULTS FROM HEARABLE DATA AND PPG

3.1 Introduction

One of the longterm goals of the research project is to assess whether a hearable can be used on its own for longitudinal monitoring of biosignals to promote early detection of AD. As part of a proof of concept, this study included research-approved sensors that provide the ground truth data for the biosignals extracted from inside the ear. One of these sensors is the Shimmer GSR+ (Shimmer, Dublin, Ireland) that contained a PPG earlobe clip. The PPG measures the blood volume from which the heart rate (HR) and HRV can be computed. These biosignals are indicators of stress and the status of the autonomic nervous system. Stress is an interesting aspect to study in the identification of AD. AD patients might feel high levels of stress where they could exhibit an increase in HR or they might let go during a difficult task and give up, leading to a low variation in biosignals or even a decline.

During the data collection, unfortunately, the Shimmer GSR+ sensor broke after the fourth participant, meaning the PPG data is only available for four CI participants. Repairing the sensor or getting a new one involved communicating with the manufacturer and shipping delays. Given the challenges associated with recruiting participants from an aging population and time constraints, we decided to proceed with data collection rather than cancel their participation. Therefore, the initial idea behind the analysis shifted from identifying AD patients to observing whether there is a correlation between the performance on the four CATs and the stress levels of the participants. Nonetheless, four participants with a reliable HR ground truth are a step forward towards a preliminary evaluation of HR taken from IEM data that is around 45 minute recording per participant.

This chapter presents the results of HR and HRV extracted from IEM and PPG from four CI participants. This chapter presents two main sub-objectives:

1. Compare the HR extracted from IEM data to the ground truth taken from PPG. This sub-objective addresses the potential of extracting clear and valid heartbeats from IEM data.
2. Check whether there is a correlation between the HR, the HRV features, and the performance of the four participants on the CATs.

3.2 Methodology

Tempbeat (Benesch *et al.*, 2024), a template-based interbeat interval extraction algorithm from in-ear biosignals, was used to extract heartbeats from IEM data. R-R interval (RRI), the intervals between consecutive R-peaks, were computed, in addition to the HR in beats per minute (bpm). The algorithm filters the in-ear signal at the start using a bandpass Burtterworth filter with a low cut-off frequency of 0.5 Hz and a high cut-off frequency at 25 Hz, making sure speech and other high frequencies are not considered in the computations. To evaluate the results of the heartbeat extraction from in-ear signals data through Tempbeat, RRI and HR were also computed from the PPG signal using the NeuroKit2 Python package (Makowski *et al.*, 2021). The mean absolute error (MAE) was computed for the RRI and HR for each of the four CI participants.

The performance on the HINT and TDT is expressed as the average SNR of both of the noise settings, babble and speech-shaped noise. The higher the SNR, the poorer the performance. As for DDT and DSI, the number of successfully identified presentations in both ears is computed and designated as the performance metric.

The HRV features are extracted using NeuroKit2 package (Makowski *et al.*, 2021) that computes the time-domain, frequency-domain, and non-linear HRV features. Three HRV features were chosen to be studied in this work. The ratio low-frequency to high-frequency (LFHF) power where parasympathetic nervous system (PNS) and SNS contribute to the LF power component and the PNS contributes to the HF component. This means that an increase ratio indicates a decrease in the PNS (Rubio *et al.*, 2025). The second HRV feature is the root mean square of successive differences between normal heartbeats (RMSSD), which represents the variation between consecutive heartbeats and serves as the main time-domain indicator for assessing parasympathetic influences on HRV (Shaffer & Ginsberg, 2017). The third component is the standard deviation of the interbeat interval of normal sinus beats (SDNN) provides a measure of global HRV, encompassing both sympathetic and parasympathetic activity (Shaffer & Ginsberg, 2017).

3.3 Results

Table 3.1 presents the results of the MAE of the RRI and HR between the data taken from IEM data and PPG. Regarding RRI, the mean MAE is equal to 158.97 ms with a standard deviation of 15.86 ms. As for the HR, the mean MAE is equal to 15.83 bpm with standard deviation of 2.18 bpm.

Table 3.1 Mean absolute error (MAE) for R-R intervals (RRI) in ms and heart rates (HR) in bpm across participants comparing the heartbeat data extracted from PPG and IEM data

Participant	MAE_RRI (ms)	MAE_HR (bpm)
P1	139.88	14.54
P2	150.01	13.28
P3	173.99	18.25
P4	172.01	17.26
Mean	**158.97**	**15.83**
Standard Deviation	**15.86**	**2.18**

Table 3.2 Scores of the four participants on the CATS. Low values on HINT and TDT indicate better performance. Low values on DDT and DSI indicate bad performance

Participant	HINT (dB)	TDT (dB)	DDT	DSI
P1	-5.46	-14.32	13	8
P2	-2.8	-12.4	0	0
P3	-0.4	-10.96	0	0
P4	-3.86	-14	6	5

Figure 3.1 and Figure 3.2 present the HR and HRV features, respectively, across all four CATs for each participant, while Table 3.2 presents the scores on the tests. For participant 1, LFHF ratio is 2.8 for HINT and TDT, 1.8 for DDT, and 1.4 for DSI. As for RMSSD, the values across the four test centers around 30 ms. SDNN is equal to 35 ms for DDT, 30 ms for HINT and TDT, and

27 ms for DSI. For participant 2, LFHF ratio is 1.2 for DSI, 1 for HINT, 0.7 for TDT, and 0.5 for DDT. The RMSSD is equal to 239 ms for HINT, 183 ms for DDT, 180 ms for DSI, and 137 ms for TDT. As for SDNN, it has a similar pattern to RMSSD where it has a high value of 168 for HINT, followed by 134, 124, and 94 ms for DSI, DDT, and TDT, respectively. Participant 3 has a LFHF ratio of 6 and 5.2 for DDT and DSI, respectively, while 3.5 for TDT and 1.2 for HINT. Highest RMSSD value is 43 ms for HINT while the other tests follow with 36 ms for TDT, and 32 ms for both DDT and DSI. As for SDNN, it varies between 48 ms for DDT, 47 ms for HINT, 45 ms for DSI, and 43 ms for TDT. Finally, regarding participant 4, all four LFHF ratios are below one with 0.9 for DSI, 0.5 for DDT, followed by 0.4 for TDT and HINT. RMSSD is equal to 158 ms for HINT, 113 ms for TDT, 105 and 103 ms for DSI and DDT, respectively. SDNN has a similar pattern with 126 ms for HINT, 85 ms for TDT, 80 ms for DSI, and 74 ms for DDT.

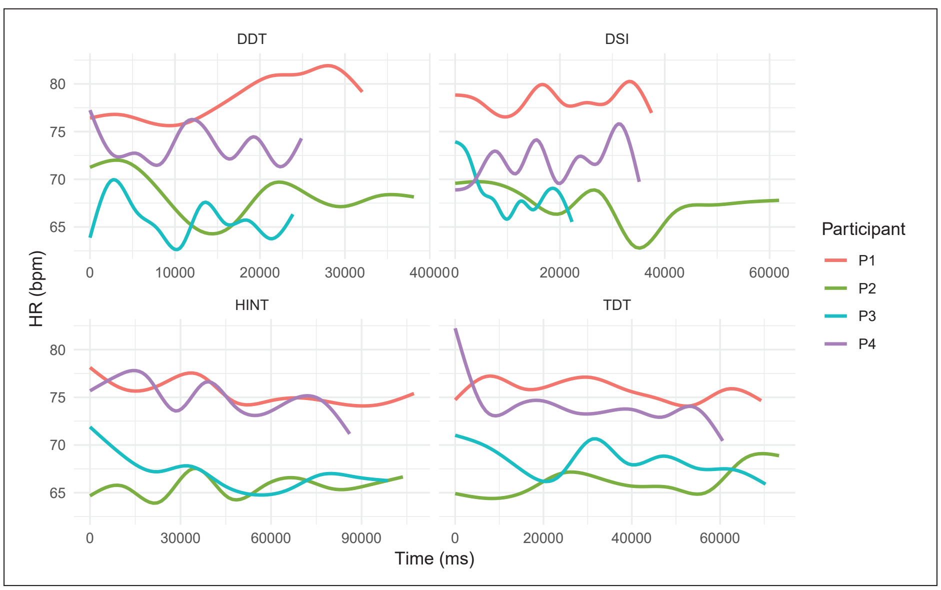


Figure 3.1 Heart rate (HR) in bpm for each CAT in ms grouped by participant

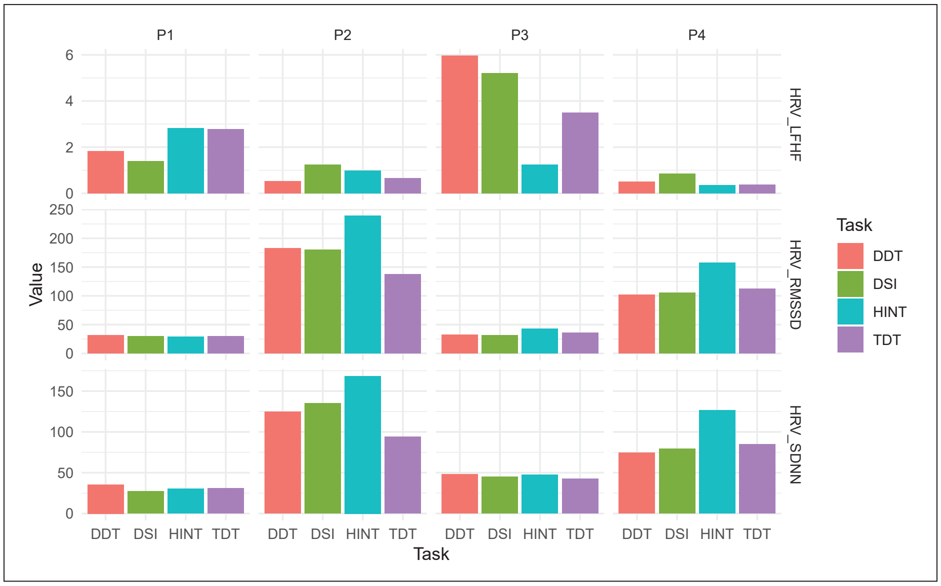
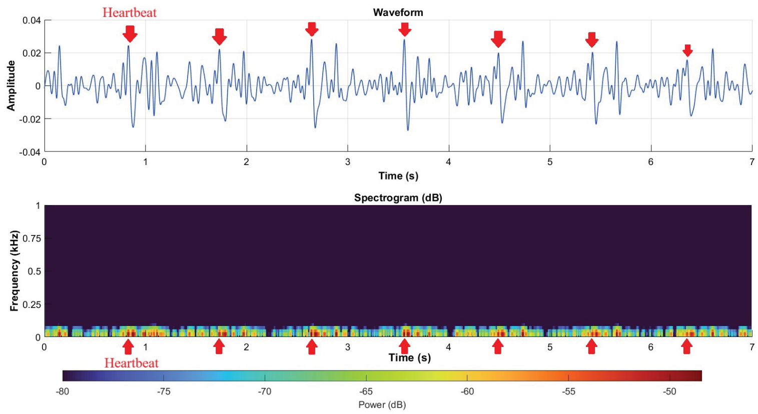


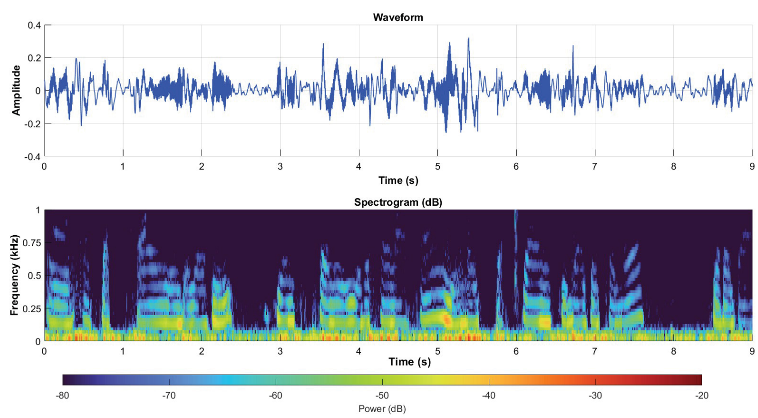
Figure 3.2 Heart rate variability (HRV) metrics for each CAT grouped by participant

3.4 Discussion

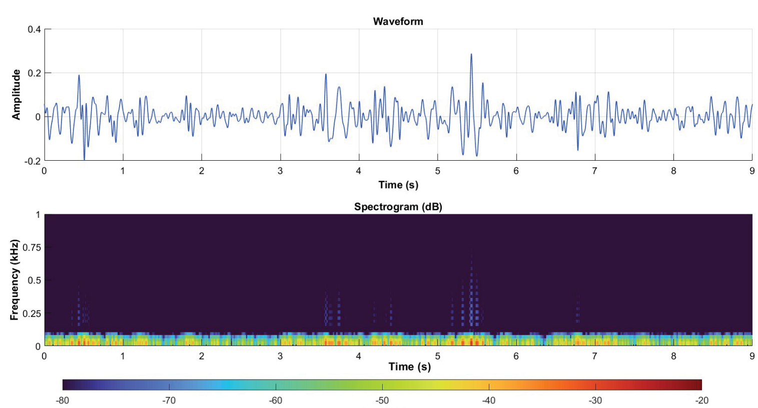
The extraction of HR and RRI from IEM data using Tempbeat resulted in a high MAE when compared to the HR and RRI extracted from an earlobe PPG sensor. After further observation, Tempbeat is not able to correctly detect heartbeats during speech, despite filtering speech before heartbeat extraction. One of the reasons might be that jaw movement during speech causes artifacts that are still present after filtering since they are very low-frequency in nature. Figure 3.3 shows a filtered waveform and spectrogram of a speech sample following the frequencies Tempbeat uses. The jaw movements artifacts that can be seen in Figure 3.3c disturb the hearbeats and interferes with the performance of Tempbeat.



a) Waveform and spectrogram of a recording without speech



b) Waveform and spectrogram of a recording with speech



c) Waveform and spectrogram of a recording with speech that was filtered

Figure 3.3 Waveforms and spectrograms showing the jaw movements artifacts after filtering. The red lines in subplot a indicate heartbeats

HR and HRV are commonly used to assess autonomic nervous system activity. HR increases indicating higher arousal or cognitive load, while changes in HRV features are related to the balance between the SNS and the PNS. The HR and HRV observed across participants, coupled with their performances on the CATs, highlight that the patterns depend on each individual. However, the resulted values appear to have been contaminated by artifacts due to general movement. The HR, HRV, and IEM data highlight the limitations of the work done in this chapter. First, the limited sample size coupled with the sensor malfunctioning put constraints on the analysis that could be done. Second, the data collection protocol did not contain a rest period during which a baseline for the biosignals for each participant can be identified. Furthermore, the use of a better HR sensor is recommended since PPG measures the HR indirectly through blood volume, in addition to its placement on the earlobe which is not at the same place of the hearable, i.e. inside the ear. However, Ferlini *et al.* (2021) show that the best placement of an ear PPG is inside the ear, which advocates for the multifunctionality of a hearable. Moreover, an algorithm that accounts for speech and jaw movements while extracting in-ear biosignals is needed for better results and cleaner data. Finally, there is a big discussion around interpreting HRV features and their relation to the autonomic nervous system and the balance of the SNS and the PNS. Hence, reviewing the choice of the HRV metrics is highly recommended.

CONCLUSION AND RECOMMENDATIONS

4.1 Conclusions

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of dementia. Early detection is essential for timely clinical intervention and lifestyle changes, yet current diagnostic methods are costly and invasive. Therefore, the need for accessible and non-invasive diagnostic methods is on the rise. Research shows that AD affects the central auditory processing system, with patients having difficulties discerning speech in noise settings or with competing messages. Other biosignals such as pupil dilation, eye movements, and speech features, could be possible biomarkers for early detection of AD. Hearables, intra-aural wearables equipped with in-ear microphones (IEMs), can capture these signals making them promising monitoring tools due to their integration in every day life. However, literature lacks in multimodal IEM data taken from AD patients, hence, a database combining these signals is needed for the progression of this research.

This work presents a multimodal database from which it explores the use of central auditory tests (CATs), particularly speech in noise and dichotic tests, as accessible and non-invasive tools for detecting cognitively impaired (CI) participants, while being administered by a hearable. Statistical analysis confirmed that CI participants performed significantly worse on CATs than cognitively unimpaired (CU) participants. Logistic regression results indicate that dichotic tests had the most weight on the predictability of CI. As the results corroborate previous literature, they also validate the pipeline followed for the administration of CATs through a hearable. Preliminary heart rate (HR) and heart rate variability (HRV) results revealed individual differences but were limited to sensor artifacts and malfunctions, small sample size, and lack of baseline. Tempbeat algorithm struggled to extract heartbeats due to jaw movements.

Overall, the findings support the administration of CATs through hearables as a practical, scalable method for identifying cognitive impairment, while highlighting technical and methodological limitations that guide future research on integrating multimodal biosignal monitoring for AD detection.

4.2 Contributions

4.2.1 Gaze and Auditory Response Database for Evaluating Neurocognitive Impairment and Alzheimer's disease (GARDENIA)

A major outcome of this work is the development of GARDENIA, a comprehensive multimodal database designed for the study of AD and mild cognitive impairment (MCI). GARDENIA includes multiple data sources, such as in-ear biosignals, eye-tracking, speech, CATs results, and picture description tasks. The database addresses the scarcity of multimodal, synchronized datasets for AD research and provides a proof of concept for future studies on sensor-based detection of neurodegeneratives disorders.

4.2.2 Central auditory tests using a hearable

This work introduces innovative approach for conducting CATs using a hearable device. The system leverages in-ear microphones and in-ear speakers to present auditory stimuli and simultaneously record physiological responses. The successful administration of the tests highlights the capability to perform these tests in uncontrolled environments and get in-the-field measurements due to the non-invasiveness, accessibility, and scalability of hearables. Additionally, this contribution adds one more functionality to hearables on top of their use as hearing aid, music mediums, and noise canceling.

4.2.3 Graphical user interface (GUI)

A custom GUI was designed and used in the data collection to administer and navigate the CATs including hearing in noise test (HINT), triple digits test (TDT), dichotic digits test (DDT), and dichotic sentence identification test (DSI) test. The interface provides an integrated environment for tests and synchronization of data coming from different sensors. It is available for researchers on the RHAD Lab Git page.

4.3 Future work and recommendations

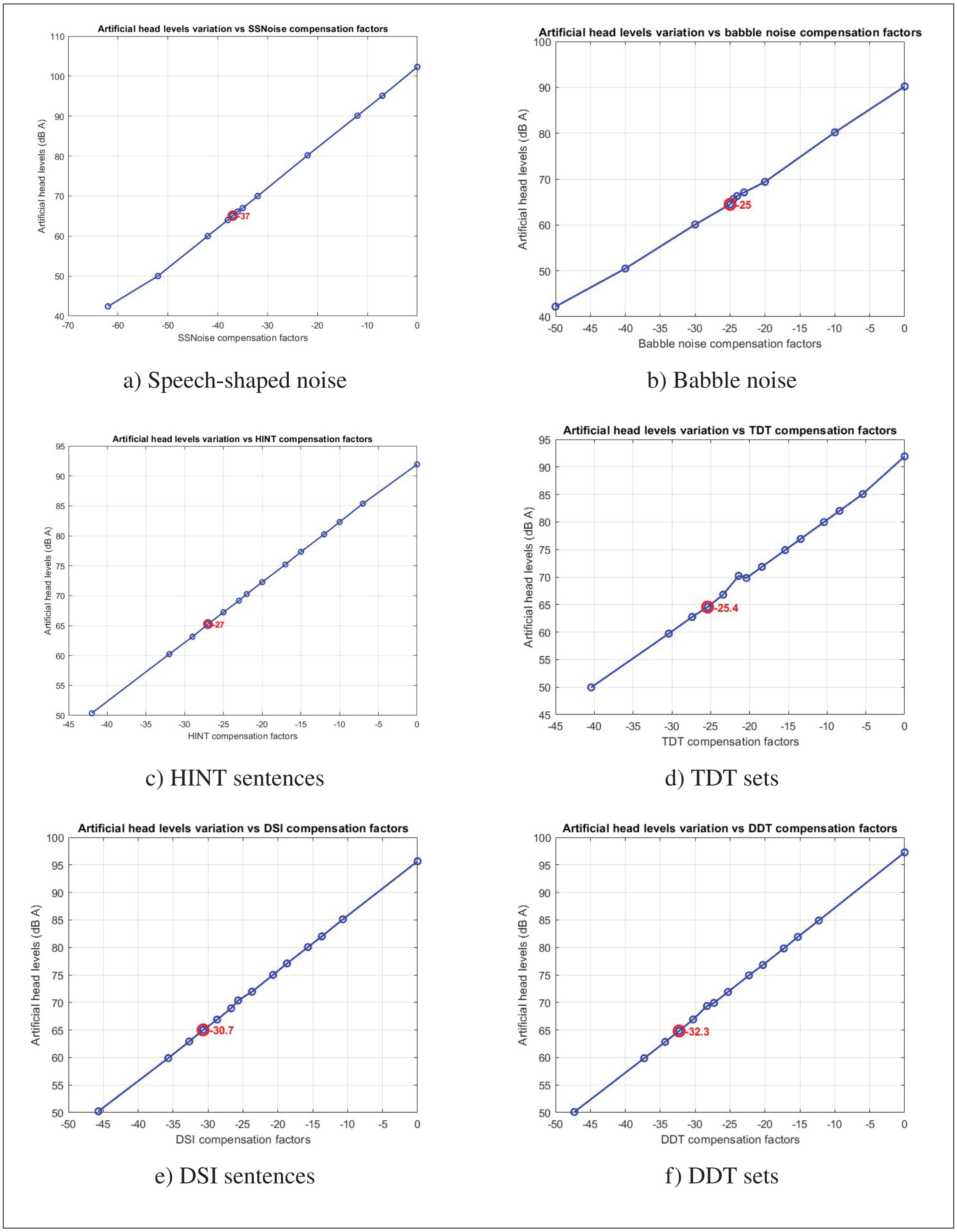
While this exploratory work forms a proof of concept for using a hearable to administer CATs and capture biosignals to detect AD, future work is needed to enhance its robustness, generalizability, and clinical relevance. Future work can tackle the following:

- Increasing the number of participants to increase the power of the statistical analysis and the range of cognitive impairment to account for generalizability.
- Having CU and CI participants age matched to eliminate the effect of age without needing to circumvent the issue by computing adjusted scores.
- Including a repeated rest period form which the baseline can be computed for each participant in order to facilitate the individual analysis of the participants.
- Having a more robust sensor that measures HR.
- Adapting Tempbeat to account for jaw movements and low-frequency artifacts that hinder the extraction of heartbeats.
- Analyzing the data related to eyes such as pupil diameters and eye movements captured by the ground truth sensor in order to identify CI participants, in addition to working on extracting these data using IEMs.
- Analyzing the picture description task data in terms of speech and lexical contents to identify CI participants.

APPENDIX I

CALIBRATION PLOTS AND DETAILED SCORES - CHAPTER 2

Calibration of the presentation level of the hearable



a) Speech-shaped noise

b) Babble noise

c) HINT sentences

d) TDT sets

e) DSI sentences

f) DDT sets

Figure A I.1 Plots showing the variation of the presentation level of different stimuli (noise and speech) presented by the hearable at the artificial head as the compensation (calibration) factor increases. The calibration factor at 65 dBA is highlighted in red

Detailed central auditory tests scores

Table A I.1 The means and standard deviations of the raw, z, predicted, adjusted, and adjusted z scores of the CI and CU groups

		CI	CU
HINT SSNoise	Raw	-1.41 (1.40)	-2.7 (0.88)
	Z-scores	-1.45 (1.59)	-8.33×10^{-11} (1)
	Predicted	-2.44 (0.14)	-2.7 (0.11)
	Adjusted	-1.02 (1.41)	-1.66×10^{-10} (0.87)
	Adjusted z-scores	-1.17 (1.61)	-1.66×10^{-10} (1)
HINT Babble	Raw	-3.9 (2.96)	-6.42 (1.33)
	Z-scores	-1.89 (2.22)	-8.33×10^{-11} (1)
	Predicted	-5.69 (0.41)	-6.42 (0.33)
	Adjusted	-1.79 (3.02)	8.33×10^{11} (1.28)
	Adjusted z-scores	-1.39 (2.34)	-8.33×10^{11} (1)
TDT SSNoise	Raw	-11.28 (2.21)	-13.41 (0.59)
	Z-scores	-3.59 (3.72)	-2.32×10^{-18} (1)
	Predicted	-13.19 (0.12)	-13.41 (0.10)
	Adjusted	-1.91 (2.19)	-3.33×10^{-10} (0.58)
	Adjusted z-scores	-3.27 (3.76)	-1.04×10^{-17} (1)
TDT Babble	Raw	-15.2 (2.99)	-17.91 (0.87)
	Z-scores	-3.10 (3.43)	5.80×10^{-18} (1)
	Predicted	-17.52 (0.22)	-17.91 (0.18)
	Adjusted	-2.32 (3.03)	8.33×10^{-11} (0.85)
	Adjusted z-scores	-2.71 (3.55)	-8.33×10^{-11} (1)

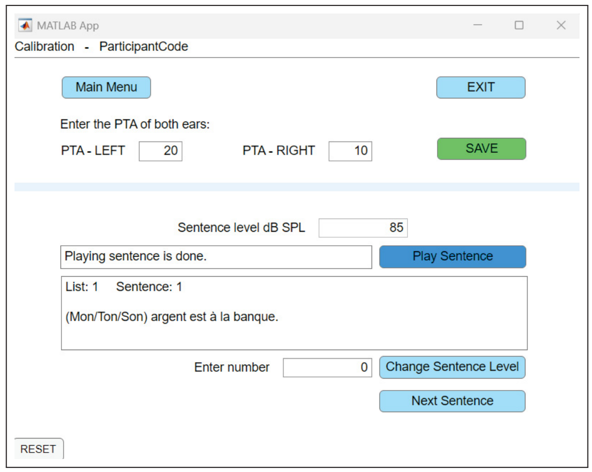
Table A I.1 The means and standard deviations of the raw, z, predicted, adjusted, and adjusted z scores of the CI and CU groups (continued)

		CI	**CU**
DDT	Raw	4.87 (5.93)	17.41 (4.20)
	Z-scores	-2.97 (1.41)	-2.5×10^{-10} (1)
	Predicted	16.16 (0.72)	17.41 (0.58)
	Adjusted	-11.28 (6.04)	-6.70×10^{-17} (4.16)
	Adjusted z-scores	-2.70 (1.45)	8.33×10^{-11} (1)
DSI	Raw	5.62 (6.78)	13.41 (4.33)
	Z-scores	-1.79 (1.56)	-8.33×10^{-11} (1)
	Predicted	11.9 (0.85)	13.41 (0.69)
	Adjusted	-6.30 (6.53)	-1.66×10^{-10} (4.28)
	Adjusted z-scores	-1.47 (1.52)	-2.30×10^{-18} (1)
Speech in SSNoise (HINT + TDT)	Z-scores	-2.527 (2.48)	1.66×10^{-10} (0.76)
	Adjusted z-scores	-2.22 (2.50)	-1.66×10^{-10} (0.75)
Speech in Babble noise (HINT + TDT)	Z-scores	-2.49 (2.61)	8.33×10^{-11} (0.83)
	Adjusted z-scores	-2.05 (2.73)	8.94×10^{-18} (0.82)
Speech in noise	Z-scores	-2.51 (2.21)	8.33×10^{-11} (0.69)
	Adjusted z-scores	-2.14 (2.27)	-4.62×10^{-18} (0.68)
Dichotics (DDT + DSI)	Z-scores	-2.38 (1.39)	-1.66×10^{-10} (0.78)
	Adjusted z-scores	-2.09 (1.39)	-8.33×10^{-11} (0.78)
CAT (HINT + TDT + DDT + DSI)	Z-scores	-2.45 (1.74)	-1.73×10^{-18} (0.65)
	Adjusted z-scores	-2.11 (1.78)	-8.33×10^{-11} (0.63)

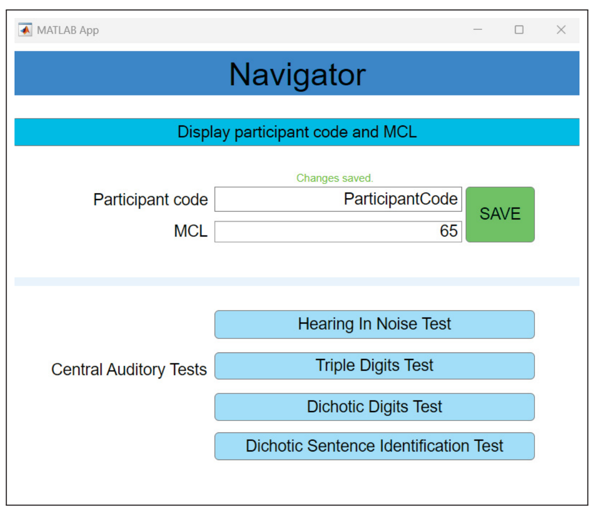
ANNEX A

GRAPHICAL USER INTERFACE

Finding the most comfortable level and navigating the tests



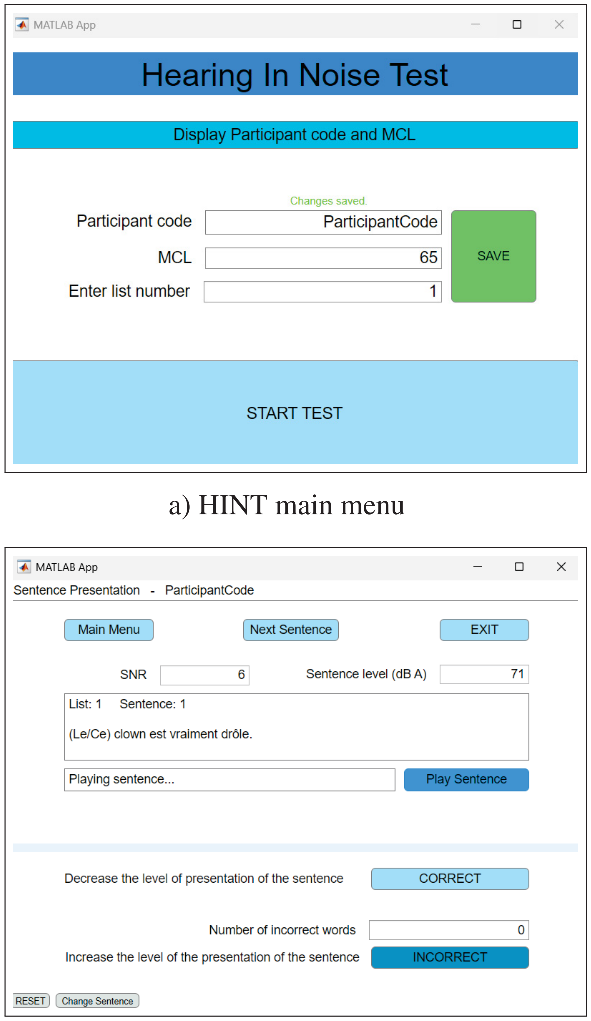
a) Interface for finding the most comfortable level



b) The navigation interface

Figure A A.1 Graphical user interface (GUI) of the data collection: (a) finding the most comfortable level and (b) navigation menu

HINT

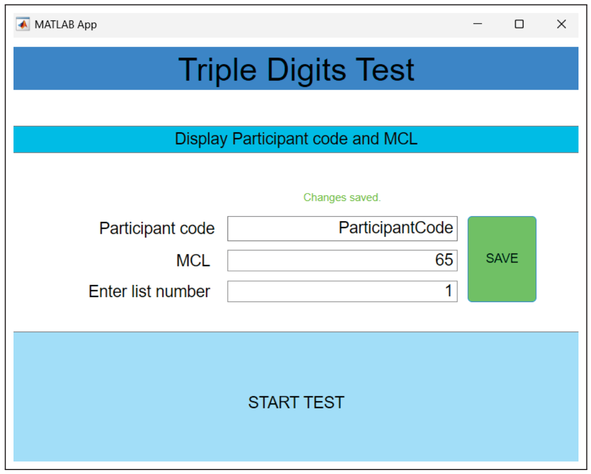


a) HINT main menu

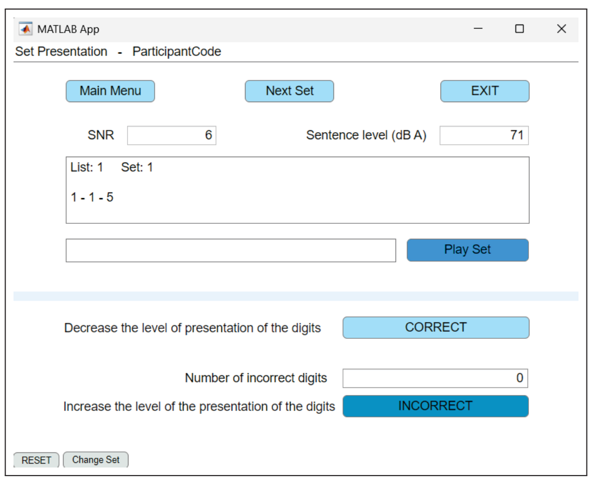
b) Interface for HINT

Figure A A.2 Graphical user interface (GUI) of the HINT test: (a) main menu and (b) testing interface

TDT



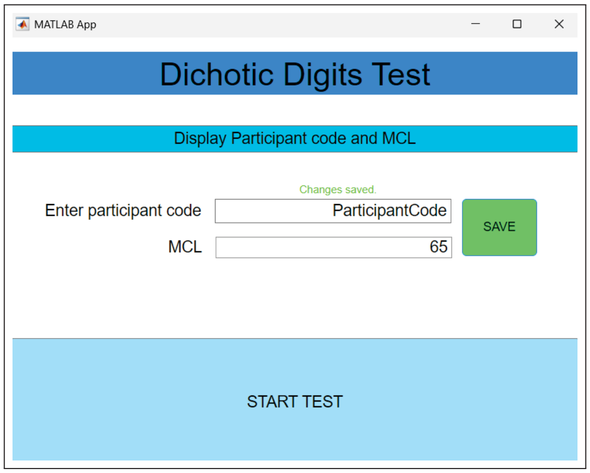
a) TDT main menu



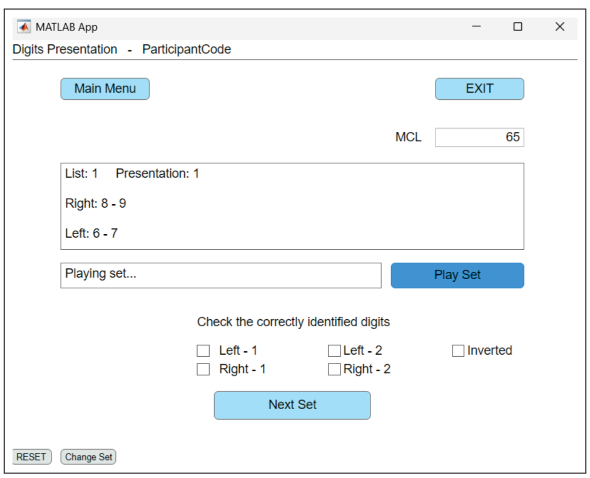
b) Interface for TDT

Figure A A.3 Graphical user interface (GUI) of the TDT test: (a) main menu and (b) testing interface

DDT



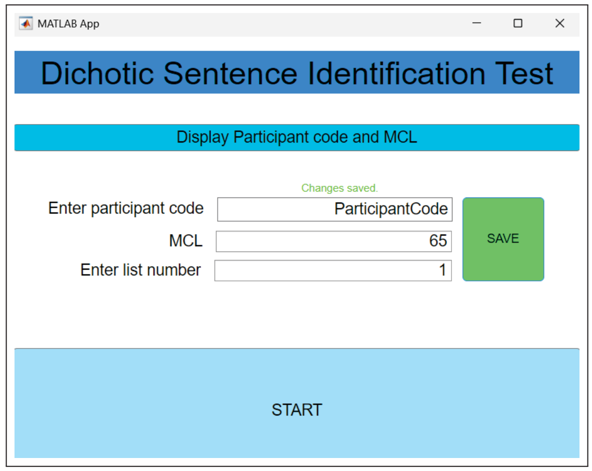
a) DDT main menu



b) Interface for DDT

Figure A A.4 Graphical user interface (GUI) of the DDT test: (a) main menu and (b) testing interface

DSI



a) DSI main menu

b) Interface for DSI

Figure A A.5 Graphical user interface (GUI) of the DSI test: (a) main menu and (b) testing interface

ANNEX B

THE PICTURE DESCRIPTION TASK

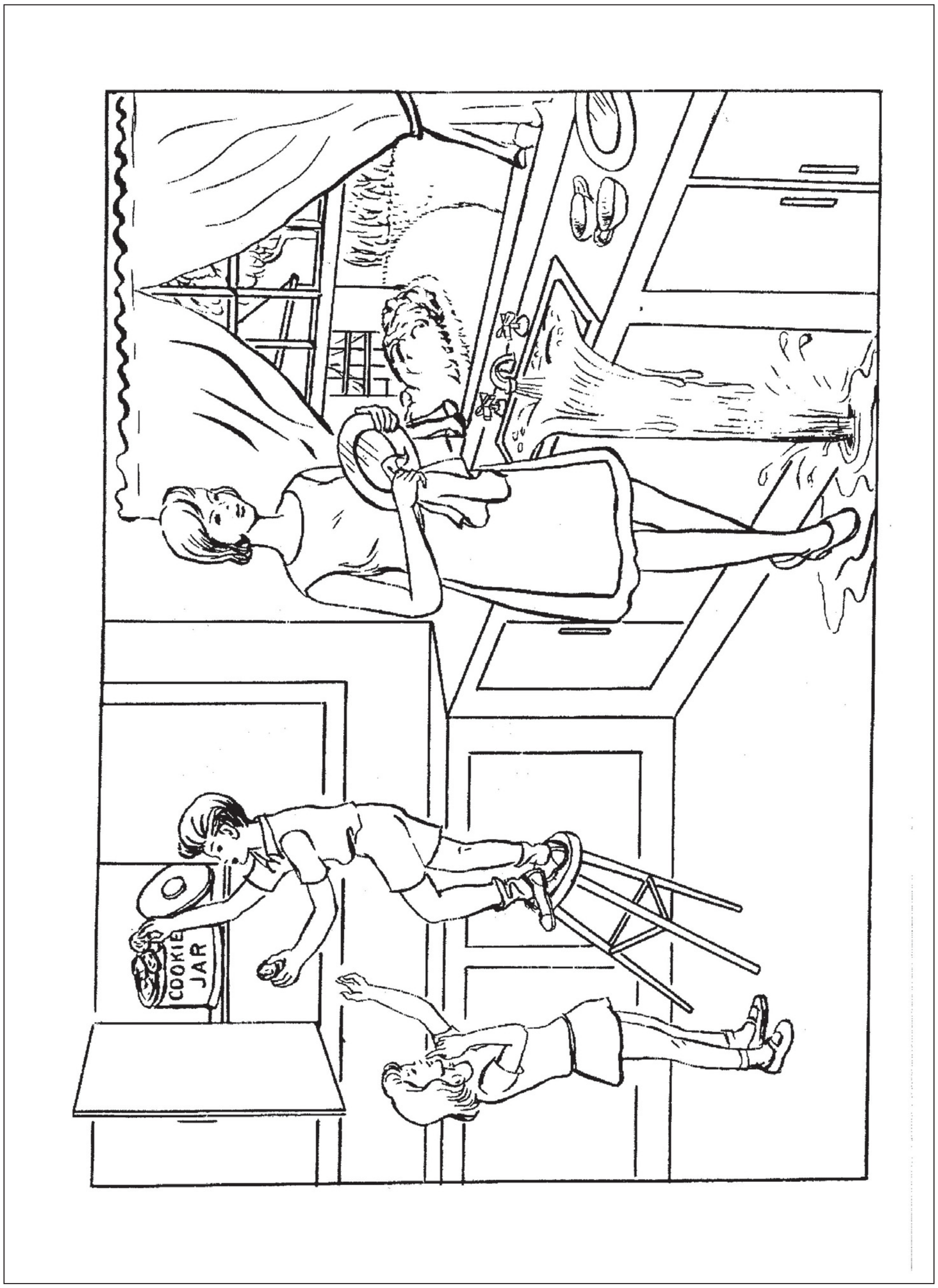


Figure A B.1 Boston Cookie Theft Picture, the simple picture that was used in the picture description task
Taken and adapted from Harold Goodglass *et al.* (2021)

BIBLIOGRAPHY

Ahmed, S., Haigh, A.-M. F., de Jager, C. A. & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, 136(12), 3727–3737.

Al Haddad, R., Chamoun, M., Tardif, C. L., Guimond, S., Horga, G., Rosa-Neto, P. & Cassidy, C. M. (2023). Normative Values of Neuromelanin-Sensitive MRI Signal in Older Adults Obtained Using a Turbo Spin Echo Sequence. *Journal of Magnetic Resonance Imaging*, 58(1), 294–300.

Alkenani, A. H., Li, Y., Xu, Y. & Zhang, Q. (2021). Predicting Alzheimer's Disease from Spoken and Written Language Using Fusion-Based Stacked Generalization. *Journal of Biomedical Informatics*, 118, 103803. doi: 10.1016/j.jbi.2021.103803.

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., Van De Pavert, S. H. P. & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, 14(4), 1–1. doi: 10.1167/14.4.1. Publisher: Association for Research in Vision and Ophthalmology (ARVO).

Alzheimer Society of Canada. (2023). *Navigating the Path Forward for Dementia in Canada: The Landmark Study Report 1*. Toronto, ON. Retrieved from: https://alzheimer.ca/landmarkstudy.

Alzheimer Society of Canada. [Accessed: 2025-09-21]. (2024). The Many Faces of Dementia in Canada: The Landmark Study Report #2. Retrieved from: https://alzheimer.ca/en/the-many-faces-of-dementia-in-canada-landmark-study-volume-2.

Alzheimer's Association. (2025a). *Alzheimer's Disease Facts and Figures*. Retrieved from: https://www.alz.org/getmedia/ef8f48f9-ad36-48ea-87f9-b74034635c1e/alzheimers-facts-and-figures.pdf.

Alzheimer's Association. (2025b). *Inside the brain: A tour of how the mind works - Brain Tour Part 2*. Retrieved from: https://www.alz.org/alzheimers-dementia/what-is-alzheimers/brain-tour-part-2.

Alzheimer's Disease International. (2022). *World Alzheimer Report 2022 – Life after diagnosis: Navigating treatment, care and support*.

Anderson, T. J. & MacAskill, M. R. (2013). Eye movements in patients with neurodegenerative disorders. *Nature Reviews Neurology*, 9(2), 74–85. doi: 10.1038/nrneurol.2012.273.

Ardila, A., Bernal, B. & Rosselli, M. (2016). The role of Wernicke's area in language comprehension. *Psychology & Neuroscience*, 9(3), 340.

Benesch, D., Chabot, P., Tom, A., Voix, J. & Bouserhal, R. E. (2024). Template-Based Extraction of Interbeat Intervals from in-Ear Heartbeat Sounds. *2024 IEEE 20th International Conference on Body Sensor Networks (BSN)*, pp. 1–4. doi: 10.1109/BSN63547.2024.10780689.

Benesch, D., Villatte, B., Vinet, A., Hébert, S., Voix, J. & Bouserhal, R. E. (2025). Stress classification with in-ear heartbeat sounds. *Computers in Biology and Medicine*, 186, 109555. doi: 10.1016/j.compbiomed.2024.109555.

Berg, L. (1984). Clinical dementia rating. *The British Journal of Psychiatry*, 145(3), 339–339.

Billings, C. J., Olsen, T. M., Charney, L., Madsen, B. M. & Holmes, C. E. (2024). Speech-in-noise testing: An introduction for audiologists. *Seminars in Hearing*, 45(01), 055–082.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M. & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of neuroscience*, 17(1), 353–362.

Bomasang-Layno, E. & Bronsther, R. (2021). Diagnosis and Treatment of Alzheimer's Disease:. *Delaware Journal of Public Health*, 7(4), 74–85. doi: 10.32481/djph.2021.09.009.

Bouma, A. & Gootjes, L. (2011). Effects of attention on dichotic listening in elderly and patients with dementia of the Alzheimer type. *Brain and Cognition*, 76(2), 286–293. doi: 10.1016/j.bandc.2011.02.008.

Bouserhal, R. E., Bernier, A. & Voix, J. (2019). An in-ear speech database in varying conditions of the audio-phonation loop. *The Journal of the Acoustical Society of America*, 145(2), 1069–1077. doi: 10.1121/1.5091777.

Butkow, K.-J., Dang, T., Ferlini, A., Ma, D. & Mascolo, C. (2023). hEARt: Motion-resilient Heart Rate Monitoring with In-ear Microphones. *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 200-209. doi: 10.1109/PERCOM56429.2023.10099317.

Chabot, P., Bouserhal, R. E., Cardinal, P. & Voix, J. (2021). Detection and classification of human-produced nonverbal audio events. *Applied Acoustics*, 171, 107643. doi: 10.1016/j.apacoust.2020.107643.

Chadwick Boulay & Tristan Stenner. (2025). *Lab Streaming Layer*. Retrieved from: https://github.com/sccn/labstreaminglayer.

Chen, Y., Chen, T. & Hou, R. (2022). Locus coeruleus in the pathogenesis of Alzheimer's disease: A systematic review. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 8(1). doi: 10.1002/trc2.12257. Publisher: Wiley.

Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., Thom, V. M., Nash, K. J., Elhamoui, H., Milligan, R. et al. (2016). Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database of Systematic Reviews*, (1).

Dautzenberg, G., Lijmer, J. & Beekman, A. (2020). Diagnostic accuracy of the Montreal Cognitive Assessment (MoCA) for cognitive screening in old age psychiatry: Determining cutoff scores in clinical practice. Avoiding spectrum bias caused by healthy controls. *International journal of geriatric psychiatry*, 35(3), 261–269.

DeTure, M. A. & Dickson, D. W. (2019). The neuropathological diagnosis of Alzheimer's disease. *Molecular neurodegeneration*, 14(1), 32.

Dickerson, B. C. & Sperling, R. A. (2008). Functional abnormalities of the medial temporal lobe memory system in mild cognitive impairment and Alzheimer's disease: insights from functional MRI studies. *Neuropsychologia*, 46(6), 1624–1635.

Doody, R. S., Thomas, R. G., Farlow, M., Iwatsubo, T., Vellas, B., Joffe, S., Kieburtz, K., Raman, R., Sun, X., Aisen, P. S. et al. (2014). Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *New England Journal of Medicine*, 370(4), 311–321.

Dunn, J., Runge, R. & Snyder, M. (2018). Wearables and the medical revolution. *Personalized medicine*, 15(5), 429–448.

Felix, R. A., Gourévitch, B. & Portfors, C. V. (2018). Subcortical pathways: Towards a better understanding of auditory disorders. *Hearing Research*, 362, 48–60. doi: 10.1016/j.heares.2018.01.008. Publisher: Elsevier BV.

Ferlini, A., Montanari, A., Min, C., Li, H., Sassi, U. & Kawsar, F. (2021). In-ear PPG for vital signs. *IEEE Pervasive Computing*, 21(1), 65–74.

Fletcher, W. A. & Sharpe, J. A. (1986). Saccadic eye movement dysfunction in Alzheimer's disease. *Annals of Neurology*, 20(4), 464–471. doi: 10.1002/ana.410200405.

Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T. & Crone, N. E. (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences*, 112(9), 2871–2875.

Folstein, M. F., Robins, L. N. & Helzer, J. E. (1983). The mini-mental state examination. *Archives of general psychiatry*, 40(7), 812–812.

Forbes-McKay, K. E. & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurological Sciences*, 26(4), 243–254. doi: 10.1007/s10072-005-0467-9.

Forbes-McKay, K., Shanks, M. F. & Venneri, A. (2013). Profiling spontaneous speech decline in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica*, 25(6), 320–327. doi: 10.1017/neu.2013.16.

Fraser, K. C., Meltzer, J. A. & Rudzicz, F. (2015). Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *Journal of Alzheimer's Disease*, 49(2), 407–422. doi: 10.3233/jad-150520. Publisher: SAGE Publications.

Gates, G. A., Beiser, A., Rees, T. S., D'Agostino, R. B. & Wolf, P. A. (2002). Central Auditory Dysfunction May Precede the Onset of Clinical Dementia in People with Probable Alzheimer's Disease. *Journal of the American Geriatrics Society*, 50(3), 482–488. doi: 10.1046/j.1532-5415.2002.50114.x.

Gates, G. A., Anderson, M. L., Feeney, M. P., McCurry, S. M. & Larson, E. B. (2008). Central Auditory Dysfunction in Older Persons With Memory Impairment or Alzheimer Dementia. *Archives of Otolaryngology–Head & Neck Surgery*, 134(7), 771. doi: 10.1001/archotol.134.7.771.

Gates, G. A., Anderson, M. L., McCurry, S. M., Feeney, M. P. & Larson, E. B. (2011). Central Auditory Dysfunction as a Harbinger of Alzheimer Dementia. *Archives of Otolaryngology–Head & Neck Surgery*, 137(4), 390. doi: 10.1001/archoto.2011.28. Publisher: American Medical Association (AMA).

Gootjes, L., Bouma, A., Van Strien, J., Schijndel, R. V., Barkhof, F. & Scheltens, P. (2006). Corpus callosum size correlates with asymmetric performance on a dichotic listening task in healthy aging but not in Alzheimer's disease. *Neuropsychologia*, 44(2), 208–217. doi: 10.1016/j.neuropsychologia.2005.05.002.

Granholm, E. L., Panizzon, M. S., Elman, J. A., Jak, A. J., Hauger, R. L., Bondi, M. W., Lyons, M. J., Franz, C. E. & Kremen, W. S. (2017). Pupillary Responses as a Biomarker of Early Risk for Alzheimer's Disease. *Journal of Alzheimer's Disease*, 56(4), 1419–1428. doi: 10.3233/jad-161078. Publisher: SAGE Publications.

Gruters, K. G., Murphy, D. L. K., Jenson, C. D., Smith, D. W., Shera, C. A. & Groh, J. M. (2018). The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. *Proceedings of the National Academy of Sciences*, 115(6). doi: 10.1073/pnas.1717948115.

Harasty, J. A., Halliday, G. M., Kril, J. & Code, C. (1999). Specific temporoparietal gyral atrophy reflects the pattern of language dissolution in Alzheimer's disease. *Brain*, 122(4), 675–686.

Harold Goodglass, Edith Kaplan & Barbara Barresi. *The assessment of aphasia and related disorders.* (ed. 3, revised). Lippincott Williams & Wilkins, 2001.

Heidarzadeh, N. & Ratté, S. (2023). 'Eye-Tracking'with words for Alzheimer's Disease detection: Time alignment of words enunciation with image regions during image description tasks. *Journal of Alzheimer's Disease*, 95(3), 855–868.

Heiko Braak, Dietmar R Thal, Estifanos Ghebremedhin & Kelly Del Tredici. (2011). Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. *Journal of Neuropathology & Experimental Neurology*, 70(11), 960–969. doi: 10.1097/NEN.0b013e318232a379.

Hernández-Domínguez, L., Ratté, S., Sierra-Martínez, G. & Roche-Bergua, A. (2018). Computer-based evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture description task. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10, 260–268.

Hernández Domínguez, L. E. (2019). *Computer-based characterization of language alterations throughout the Alzheimer's disease continuum.* (Ph.D. thesis, École de technologie supérieure).

Holmes, D. (2014). Bruno Dubois: transforming the diagnosis of Alzheimer's disease. *The Lancet Neurology*, 13(6), 541. doi: 10.1016/S1474-4422(14)70106-1.

Iannitelli, A. F., Kelberman, M. A., Lustberg, D. J., Korukonda, A., McCann, K. E., Mulvey, B., Segal, A., Liles, L. C., Sloan, S. A., Dougherty, J. D. et al. (2023). The neurotoxin DSP-4 dysregulates the locus coeruleus-norepinephrine system and recapitulates molecular and behavioral aspects of prodromal neurodegenerative disease. *Eneuro*, 10(1).

Jandackova, V. K., Scholes, S., Britton, A. & Steptoe, A. (2024). Midlife heart rate variability and cognitive decline: A large longitudinal cohort study. *International Journal of Clinical and Health Psychology*, 24(4), 100518. doi: 10.1016/j.ijchp.2024.100518. Publisher: Elsevier BV.

Joshua A. Waxenbaum, Vamsi Reddy & Matthew A. Varacallo. Anatomy, Autonomic Nervous System. In *Anatomy, Autonomic Nervous System*. Treasure Island (FL): StatPearls Publishing. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK539845/.

Julayanont, P., Phillips, N., Chertkow, H. & Nasreddine, Z. S. (2012). Montreal Cognitive Assessment (MoCA): concept and clinical review. In *Cognitive screening instruments: A practical approach* (pp. 111–151). Springer.

Juva, K., Sulkava, R., Erkinjuntti, T., Ylikoski, R., Valvanne, J. & Tilvis, R. (1995). Usefulness of the Clinical Dementia Rating scale in screening for dementia. *International Psychogeriatrics*, 7(1), 17–24.

Khan, T. (2016). Chapter 2 - Clinical Diagnosis of Alzheimer's Disease. In Khan, T. (Ed.), *Biomarkers in Alzheimer's Disease* (pp. 27-48). Academic Press. doi: https://doi.org/10.1016/B978-0-12-804832-0.00002-X.

King, C. D., Lovich, S. N., Murphy, D. L., Landrum, R., Kaylie, D., Shera, C. A. & Groh, J. M. (2023). Individual similarities and differences in eye-movement-related eardrum oscillations (EMREOs). *Hearing research*, 440, 108899.

Kivipelto, M., Mangialasche, F. & Ngandu, T. (2018). Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nature Reviews Neurology*, 14(11), 653–666.

Kothe, C., Shirazi, S. Y., Stenner, T., Medine, D., Boulay, C., Grivich, M. I., Artoni, F., Mullen, T., Delorme, A. & Makeig, S. (2025). The lab streaming layer for synchronized multimodal recording. *Imaging Neuroscience*, 3, IMAG.a.136. doi: 10.1162/IMAG.a.136.

König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P. H. & David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 112–124. doi: 10.1016/j.dadm.2014.11.012.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B. & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*, 25(6), 763–770. doi: 10.1037/a0024238. Publisher: American Psychological Association (APA).

Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H. N., Ferrucci, L., Simonsick, E. M. et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA internal medicine*, 173(4), 293–299.

Lipschutz, B., Kolinsky, R., Damhaut, P., Wikler, D. & Goldman, S. (2002). Attention-Dependent Changes of Activation and Connectivity in Dichotic Listening. *NeuroImage*, 17(2), 643–656. doi: 10.1006/nimg.2002.1184.

Liu, K. Y., Elliott, T., Knowles, M. & Howard, R. (2022). Heart rate variability in relation to cognition and behavior in neurodegenerative diseases: A systematic review and meta-analysis. *Ageing Research Reviews*, 73, 101539. doi: 10.1016/j.arr.2021.101539. Publisher: Elsevier BV.

Lovich, S. N., King, C. D., Murphy, D. L. K., Landrum, R. E., Shera, C. A. & Groh, J. M. (2023). Parametric information about eye movements is sent to the ears. *Proceedings of the National Academy of Sciences*, 120(48), e2303562120. doi: 10.1073/pnas.2303562120.

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C. & Chen, S. H. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4), 1689–1696. doi: 10.3758/s13428-020-01516-y.

Martin, A. & Voix, J. (2018). In-Ear Audio Wearable: Measurement of Heart and Breathing Rates for Health and Safety Monitoring. *IEEE Transactions on Biomedical Engineering*, 65(6), 1256–1263. doi: 10.1109/TBME.2017.2720463.

Mehrban, M. H. K., Voix, J. & Bouserhal, R. E. (2024). Classification of Breathing Phase and Path with In-Ear Microphones. *Sensors*, 24(20), 6679. doi: 10.3390/s24206679.

Menezes, A. S., Coelho, A., Morais-Ribeiro, R., Oliveira, T. G., Initiative, A. D. N. et al. (2025). Atrophy of the auditory cortex in Alzheimer's Disease. *Hearing Research*, 109430.

Mestach, M., Hartsuiker, R. J. & Pistono, A. (2024). Can we track the progression of Alzheimer's Disease via lexical-semantic variables in connected speech? *Journal of Neurolinguistics*, 70, 101189.

Molitor, R. J., Ko, P. C. & Ally, B. A. (2015). Eye movements in Alzheimer's disease. *Journal of Alzheimer's disease*, 44(1), 1–12.

Mosimann, U. P., Müri, R. M., Burn, D. J., Felblinger, J., O'Brien, J. T. & McKeith, I. G. (2005). Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies. *Brain*, 128(6), 1267–1276. doi: 10.1093/brain/awh484.

Mueller, K. D., Hermann, B., Mecollari, J. & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *Journal of Clinical and Experimental Neuropsychology*, 40(9), 917–939. doi: 10.1080/13803395.2018.1446513.

Musaeus, C. S., Waldemar, G., Andersen, B. B., Høgh, P., Kidmose, P., Hemmsen, M. C., Rank, M. L., Kjær, T. W. & Frederiksen, K. S. (2022). Long-term EEG monitoring in patients with Alzheimer's disease using ear-EEG: a feasibility study. *Journal of Alzheimer's Disease*, 90(4), 1713–1723.

Musiek, F. E. & Baran, J. A. (2018). *The auditory system: Anatomy, physiology, and clinical correlates*. Plural Publishing.

Nair, S. S., Govindankutty, M. M., Balakrishnan, M., Prasad, K., Sathyaprabha, T. N. & Udupa, K. (2023). Investigation of Autonomic Dysfunction in Alzheimer's Disease—A Computational Model-Based Approach. *Brain Sciences*, 13(9), 1322.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L. & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

Nasreen, S., Hough, J. & Purver, M. Detecting Alzheimer's Disease using Interactional and Acoustic features from spontaneous speech.

Nauta, W. J. & Feirtag, M. (1979). The organization of the brain. *Scientific American*, 241(3), 88–111.

Niemczak, C. E., Lichtenstein, J. D., Magohe, A., Amato, J. T., Fellows, A. M., Gui, J., Huang, M., Rieke, C. C., Massawe, E. R., Boivin, M. J., Moshi, N. & Buckey, J. C. (2021). The Relationship Between Central Auditory Tests and Neurocognitive Domains in Adults Living With HIV. *Frontiers in Neuroscience*, 15, 696513. doi: 10.3389/fnins.2021.696513.

Nilsson, M., Soli, S. D. & Sullivan, J. A. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, 95(2), 1085–1099. doi: 10.1121/1.408469.

Ortiz, K. Z., De Lira, J. O., Minett, T. S. C. & Bertolucci, P. H. F. (2024). Language impairments in Alzheimer´s disease: What changes can be found between mild and moderate stages of the disease? *Clinics*, 79, 100412. doi: 10.1016/j.clinsp.2024.100412. Publisher: Elsevier BV.

Petti, U., Baker, S. & Korhonen, A. (2020). A systematic literature review of automatic Alzheimer's disease detection from speech and language. *Journal of the American Medical Informatics Association*, 27(11), 1784–1797. doi: 10.1093/jamia/ocaa174.

Rasmussen, J. & Langerman, H. (2019). Alzheimer's disease–why we need early diagnosis. *Degenerative neurological and neuromuscular disease*, 123–130.

Roberts, R. & Knopman, D. S. (2013). Classification and Epidemiology of MCI. *Clinics in Geriatric Medicine*, 29(4), 753–772. doi: 10.1016/j.cger.2013.07.003.

Rubio, T. d. A., Godoy, M. F. d., Uyemura, J. R. R., Tanus-Santos, J. E., Minari, T. P., Vilela-Martin, J. F., Yugar-Toledo, J. C. & Moreno Junior, H. (2025). A new approach to interpreting the LF/HF ratio: a study in hypertensive patients. *Journal of Human Hypertension*, 1–8.

Röddiger, T., Clarke, C., Breitling, P., Schneegans, T., Zhao, H., Gellersen, H. & Beigl, M. (2022). Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(3), 1–57. doi: 10.1145/3550314.

Seol, H. Y. & Moon, I. J. (2022). Hearables as a gateway to hearing health care. *Clinical and Experimental Otorhinolaryngology*, 15(2), 127–134.

Shaffer, F. & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in public health*, 5, 258.

Sheppard, M., Rasgado-Toledo, J., Duncan, N., Elliott, R., Garza-Villarreal, E. A. & Muhlert, N. (2024). Noradrenergic alterations associated with early life stress. *Neuroscience & Biobehavioral Reviews*, 164, 105832. doi: 10.1016/j.neubiorev.2024.105832. Publisher: Elsevier BV.

Sibahi, A., Gandhi, R., Al-Haddad, R., Therriault, J., Pascoal, T., Chamoun, M., Boutin-Miller, K., Tardif, C., Rosa-Neto, P. & Cassidy, C. M. (2023). Characterization of an automated method to segment the human locus coeruleus. *Human Brain Mapping*, 44(9), 3913–3925.

Stanford, S. C. (2001). Noradrenaline. In *Neurotransmitters, Drugs and Brain Function* (ed. 1, pp. 163–185). Wiley. doi: 10.1002/0470846577.ch8.

Sweetow, R. W. & Pirzanski, C. Z. (2003). The occlusion effect and ampclusion effect. *Seminars in Hearing*, 24(04), 333–344.

Swords, G. M., Nguyen, L. T., Mudar, R. A. & Llano, D. A. (2018a). Auditory system dysfunction in Alzheimer disease and its prodromal states: A review. *Ageing Research Reviews*, 44, 49–59. doi: 10.1016/j.arr.2018.04.001.

Swords, G. M., Nguyen, L. T., Mudar, R. A. & Llano, D. A. (2018b). Auditory system dysfunction in Alzheimer disease and its prodromal states: A review. *Ageing Research Reviews*, 44, 49–59. doi: 10.1016/j.arr.2018.04.001.

Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J. & Pakaski, M. (2015). Speaking in Alzheimer's Disease, is That an Early Sign? Importance of Changes in Language Abilities in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 7. doi: 10.3389/fnagi.2015.00195. Publisher: Frontiers Media SA.

Tarawneh, H. Y., Menegola, H. K., Peou, A., Tarawneh, H. & Jayakody, D. M. P. (2022). Central Auditory Functions of Alzheimer's Disease and Its Preclinical Stages: A Systematic Review and Meta-Analysis. *Cells*, 11(6), 1007. doi: 10.3390/cells11061007.

Terry, R. D. (2006). Alzheimer's disease and the aging brain. *Journal of geriatric psychiatry and neurology*, 19(3), 125–128.

Thamizhmani, L., Ganapathy, K., Palaniswamy, H. P., Pitchaimuthu, A. N. & Adhikari MR, P. (2025a). Auditory processing and its cognitive correlates in older adults with mild cognitive impairment. *BMC geriatrics*, 25(1), 373.

Thamizhmani, L., Ganapathy, K., Palaniswamy, H. P., Pitchaimuthu, A. N. & Adhikari MR, P. (2025b). Auditory processing and its cognitive correlates in older adults with mild cognitive impairment. *BMC geriatrics*, 25(1), 373.

Therriault, J., Zimmer, E. R., Benedet, A. L., Pascoal, T. A., Gauthier, S. & Rosa-Neto, P. (2022). Staging of Alzheimer's disease: past, present, and future perspectives. *Trends in Molecular Medicine*, 28(9), 726–741. doi: 10.1016/j.molmed.2022.05.008. Publisher: Elsevier BV.

Thomson, R. S., Auduong, P., Miller, A. T. & Gurgel, R. K. (2017). Hearing loss as a risk factor for dementia: A systematic review. *Laryngoscope Investigative Otolaryngology*, 2(2), 69–79. doi: 10.1002/lio2.65.

Toth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatloczki, G., Banreti, Z., Pakaski, M. & Kalman, J. (2018). A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech. *Current Alzheimer Research*, 15(2), 130–138. doi: 10.2174/1567205014666171121114930. Publisher: Bentham Science Publishers Ltd.

Tuwaig, M., Savard, M., Jutras, B., Poirier, J., Collins, D. L., Rosa-Neto, P., Fontaine, D. & Breitner, J. C. (2017). Deficit in Central Auditory Processing as a Biomarker of Pre-Clinical Alzheimer's Disease. *Journal of Alzheimer's Disease*, 60(4), 1589–1600. doi: 10.3233/JAD-170545.

Vander Werff, K. R., Niemczak, C. E. & Morse, K. (2021). Informational Masking Effects of Speech Versus Nonspeech Noise on Cortical Auditory Evoked Potentials. *Journal of Speech, Language, and Hearing Research*, 64(10), 4014–4029. doi: 10.1044/2021_JSLHR-21-00048.

Wang, J., Lu, X., Sang, J., Cai, J. & Zheng, C. (2022). Effects of Stimulation Position and Frequency Band on Auditory Spatial Perception with Bilateral Bone Conduction. *Trends in Hearing*, 26, 23312165221097196. doi: 10.1177/23312165221097196.

Wankerl, S., Nöth, E. & Evert, S. (2017). An N-Gram Based Approach to the Automatic Diagnosis of Alzheimer's Disease from Spoken Language. *Interspeech*, pp. 3162–3166.

Weller, J. & Budson, A. (2018). Current understanding of Alzheimer's disease diagnosis and treatment. *F1000Research*, 7, F1000–Faculty.

Wu, J., Toporek, A., Lin, Q., Goldstein, F. C., Loring, D. W., Kelberman, M. A., Weinshenker, D., Levey, A. I., Lah, J. J. & Qiu, D. (2025). Probing locus coeruleus functional network in healthy aging and its association with Alzheimer's disease biomarkers using pupillometry. *Alzheimer's Research & Therapy*, 17(1). doi: 10.1186/s13195-025-01701-1. Publisher: Springer Science and Business Media LLC.

Xue, Y. (2019). A review on intelligent wearables: Uses and risks. *Human Behavior and Emerging Technologies*, 1(4), 287–294.

Yancheva, M. & Rudzicz, F. (2016). Vector-space topic models for detecting Alzheimer's disease. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 2337–2346.