CORTICAL CHARACTERISTICS OF LISTENING EFFORT

Mohammadi, Yousef

DOI (link to publication from Publisher):
10.54337/aau561816351

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Mohammadi, Y. (2023). *CORTICAL CHARACTERISTICS OF LISTENING EFFORT*. Aalborg Universitetsforlag. https://doi.org/10.54337/aau561816351

CORTICAL CHARACTERISTICS OF LISTENING EFFORT

BY
YOUSEF MOHAMMADI

DISSERTATION SUBMITTED 2023

AALBORG UNIVERSITY
DENMARK

CORTICAL CHARACTERISTICS OF LISTENING EFFORT

by

Yousef Mohammadi

AALBORG UNIVERSITY
DENMARK

Submitted for the degree of

Doctor of Philosophy, Biomedical Science and Engineering

Dissertation submitted:	June 2023
PhD supervisor:	Professor Ole Kæseler Andersen Aalborg University
PhD committee:	Associate Professor Romulus Lontis (chair) Aalborg University, Denmark Associate Professor Jens Hjortkær Technical University of Denmark, Denmark Professor Nathan Weisz Paris Lodron University Salzburg, Austria
PhD Series:	Faculty of Medicine, Aalborg University
Department:	Department of Health Science and Technology

ISSN (online): 2246-1302
ISBN (online): 978-87-7573-679-9

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Stibo Complete, 2023

CV

Yousef Mohammadi received his bachelor's degree in Electrical Engineering from Malayer University, Hamadan, Iran, in 2016. He obtained his master's degree in Biomedical Engineering from Amikabir University of Tehran, Tehran, Iran, in 2019. Following his education, he started his Ph.D. at the Integrative Neuroscience group at Aalborg University under the supervision of Professor Ole Kæseler Andersen in 2020. His Ph.D. project was on the cortical characteristic of listening effort. As a part of his Ph.D. program, he joined Professor Tobias Reichenbach's research group at Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Erlangen, Germany, to conduct external research from March to July 2022. His main research interest includes auditory cognitive neuroscience, speech, and language.

List of publications:

- Mohammadi Yousef, Jan Østergaard, Carina Graversen, Ole Kæseler Andersen, and José Biurrun Manresa (2023). **"Validity and reliability of self-reported and neural measures of listening effort".** *European Journal of Neuroscience*, Under review. (Authorea Preprints; 2023. DOI: 10.22541/au.167850559.95803105/v1)

- Mohammadi Yousef, Carina Graversen, Jan Østergaard, Ole Kaeseler Andersen, and Tobias Reichenbach. **"Phase-locking of neural activity to the envelope of speech in the delta frequency band reflects differences between word lists and sentences."** *Journal of Cognitive Neuroscience*, 2023.

- Mohammadi Yousef, Carina Graversen, José Biurrun Manresa, Jan Østergaard, and Ole Kaeseler Andersen. "**Effects of background noise and linguistic violations on frontal theta oscillations during effortful listening**". *Ear and Hearing*, 2023. Under review.

- Mohammadi Yousef, Mohadeseh Shafiei Kafraj, Carina Graversen, and Mohammad Hassan Moradi (2023) "**Decreased resting-state alpha self-synchronization in depressive disorder**". *Clinical EEG and Neuroscience*.

- Mohammadi Yousef, Mojtaba Hajian, and Mohammad Hassan Moradi (2019). “**Discrimination of depression levels using machine learning methods on EEG signals.”** 27th *Iranian conference on electrical engineering (ICEE)* (pp. 1765-1769). IEEE.

- Mohammadi Yousef and Mohammadi Hassan Moradi (2020)**. “Prediction of depression severity scores based on functional connectivity and complexity of the EEG signal.”** *Clinical EEG and Neuroscience*, 52(1), 52-60.

SUMMARY ENGLISH

Listening effort is a major complaint of people with hearing loss. In recent years there has been growing interest in an objective quantification of listening effort. In this thesis, we further investigated neural measures alongside self-reported measures of listening effort with normal-hearing participants performing speech-in-noise tasks. Behavioral and electroencephalography (EEG) data were collected while participants listened to regular sentences and random word lists (with no syntactic and sentences-level semantic information) in background noise at different signal-to-noise ratios (SNRs) (-9, -6, -3, 0 dB).

The thesis consists of three studies. In the first study, we investigated the effect of SNR and linguistic violations (operationalized by sentences and word lists) and their interaction on listening effort. Behavioral and EEG data showed a significant interaction of SNR and linguistic violation on the self-reported listening effort rating and frontal theta band (4 - 7 Hz) power during the memory retention interval. That is, increased theta power was found for word lists than sentences at only SNR 0 dB, representing increased memory load for word slits. Decreasing SNR showed no significant effect on frontal theta power in sentences and word lists.

In study two, we assessed the reliability of self-reported and frontal midline theta power as measures of listening effort. Generally, self-reported measures showed an acceptable between-session variability that could be interpreted as a reliable measure of listening effort. However, frontal midline theta power showed greater between-session variability, leading us to be cautious about interpreting it as a reliable measure of listening effort, despite the acceptable effect sizes shown in study one.

In study three, we examined the effect of SNR and linguistic violations and their interaction on cortical speech tracking, quantified by the phase-locking value (PLV). We then related PLV values to the self-reported listening effort. Results showed an interaction of SNR and linguistic violation on PLV values in the delta frequency band (1-4 Hz). We observed an increased PLV for sentences compared to word lists at SNR 0 dB in sentences, indicating that the delta band speech tracking reflects linguistic information consistent with the literature. PLV only in sentences was also modulated by SNR that showed high PLV at SNR 0 dB compared to -9 dB. A marginally significant negative relationship between speech tracking and self-reported effort was observed in sentences; this trend corroborates the literature showing that decreased speech tracking is associated with increased listening effort.

The findings of this work emphasize the role of linguistic information in speech perception. This is evident from studies one and three, which showed lower frontal theta power during the memory retention interval, likely representing memory load, and higher delta-band PLV at SNR 0 dB in sentences compared to word lists. The results also indicated the effect of background noise that decreased delta band PLV in sentences.

DANSK RESUME

For at følge med i samtaler, kræves der ofte en stor lytteindsats hos mennesker med et høretab.

I de senere år har der været stigende interesse for en objektiv kvantificering af lytteindsatsen. I denne afhandling undersøgte vi neurale mål sammen med selvrapporterede målinger af lytteindsatsen for normalthørende deltagere, der lytter til tale i støj. Adfærdsmæssige og elektroencefalografiske (EEG) data blev indsamlet, mens deltagerne lyttede til almindelige sætninger og tilfældige ordlister (uden syntaktisk eller semantisk information på sætningsniveau) i baggrundsstøj og ved forskellige signal-støj-forhold (-9, -6, -3, 0 dB).

Specialet består af tre studier. I det første studie undersøgte vi effekten af signal-støj-forhold samt effekten af sætninger versus ordlister (herefter kaldet sætning-ordliste-forhold) og deres vekselvirkning på lytteindsatsen. Adfærds- og EEG-data viste en signifikant vekselvirkning mellem signal-støj-forhold og sætning-ordliste-forhold i den selvrapporterede lytteindsats samt i det frontale theta-bånd (4 - 7 Hz), som blev observeret i retention-perioden, dvs. den periode hvor deltagerne skulle huske hvad der blev sagt. Dette betyder, at der blev observeret en øget theta-effekt for ordlister i forhold til sætninger ved 0 dB signal-støj-forhold, hvilket svarer til en øget belastning af hukommelsen for ordlister. Et faldende signal-til-støj-forhold viste ingen signifikant effekt på den frontale theta for sætninger og ordlister.

I det andet studie vurderede vi pålideligheden af at bruge selvrapportering og den observerede frontal-theta effekt som mål for lytteindsatsen. Generelt viste det selvrapporterede mål en acceptabel variation mellem sessionerne og det vurderes derfor som værende et pålideligt mål for lytteindsatsen. På den andet side viste frontal theta-effekten en større variation mellem sessionerne, hvilket medfører, at vi er mere forsigtige med at fortolke det som et pålideligt mål for lytteindsatsen - på trods af de acceptable værdier vist i det første studie.

I det tredje studie undersøgte vi effekten af signal-støj-forhold og sætning-ordliste-forhold samt deres vekselvirkning på den kortikale faselåsning til talestimuli, kvantificeret ved faselåsningsværdien. Vi relaterede faselåsningsværdierne til den selvrapporterede lytteindsats. Resultaterne viste en vekselvirkning mellem signal-støj-forhold og sætning-ordliste-forhold på faselåsningsværdier i delta-frekvensbåndet (1-4 Hz). Vi observerede en øget faselåsningsværdi for sætninger i forhold til ordlister ved 0 dB signal-støj-forhold. Dette viser, at den sproglige information afspejles i deltabåndets faselåsning, hvilket er i overensstemmelse med litteraturen.

Faselåsningsværdierne i sætninger blev også moduleret af signal-støj-forholdet, som viste en højere faselåsningsværdi ved 0 dB signal-støj-forhold sammenlignet med ved -9 dB. En marginalt signifikant negativ sammenhæng mellem faselåsningsværdierne og den selvrapporterede lytteindsats blev observeret i sætninger; denne tendens bekræftes af litteraturen, der viser, at en nedsat faselåsning er forbundet med en øget lytteindsats.

Resultaterne af vores studier understreger vigtigheden af den sproglige information for taleopfattelsen. Dette er tydeligt fra studie et og tre, hvor der blev påvist en lavere frontal theta-effekt i retention-perioden for sætninger sammenlignet med ordlister, hvilket sandsynligvis repræsenterer en øget hukommelsesbelastning, samt en højere faselåsningsværdi i delta-båndet ved 0 dB signal-støj-forhold. Resultaterne indikerede også at baggrundsstøj øgede belastningen på hukommelsen og reducerede deltabåndets faselåsningsværdier for sætninger.

ACKNOWLEDGMENTS

This Ph.D. thesis is the result of almost three years of intense effort, filled with tough, exciting, frustrating, and happy moments, not just for me, but for many people who contributed to it.

Firstly, I would like to thank my supervisor, Professor Ole Kæseler Anderson for his support and help during my Ph.D. study. I consider him not only a great supervisor for my Ph.D. but also a role model and person of prominence. I want to extend a special thank you to my co-supervisors, Professor Jan Østergaard for being always helpful and supportive; and Assistant Professor Carina Graversen for her technical assistance. My gratitude also goes out to Dr. José Biurrun Manresa and Professor Tobias Reichenbach for their thoughtful comments and reflections. Their critical and sharp minds always steered me on the right track and provide me with valuable knowledge.

I want to give my warmest gratitude to my wife Mohadeseh Shafiei Kafraj for her constant support, encouragement, and love. I also would like to express my outermost thanks to my parents and siblings for their immense support and love.

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 1. INTRODUCTION

This chapter provides a broad introduction to key concepts and reviews the state-of-the-art related to self-reported and neural measures of listening effort.

1.1 Introduction

Daily listening takes place in noisy environments such as crowded restaurants and bars. Following a conversation in these situations is challenging for people with hearing impairments and might even be difficult for people with normal hearing. Hearing aids could filter out ambient noise and help hearing-impaired people understand speech, but despite their use, they complain of tiredness and fatigue after listening in a noisy environment. Tiredness arises from an extra allocation of mental resources or energy to understand speech for instance in noisy situations (Kathleen Pichora-Fuller et al., 1995; Wingfield et al., 2016). This process is called *listening effort*. Listening effort is defined as "the conscious allocation of mental resources to overcome obstacles in goal pursuit when performing a listening task" (Pichora-Fuller et al., 2016). Listening effort can also broadly be illustrated as two people with the same hearing ability performing at the same level (e.g. percentage correct) on a given listening task, while one person exerts more effort than the other.

There is broad consensus that the primary cognitive resources contributing to effortful listening are working memory and selective attention (Francis & Love, 2020). Cowan et al. (2005) have defined working memory as a "set of mental processes that maintain limited information in a transiently accessible state in the service of cognition". We use working memory to maintain relevant information and inhibit or suppress irrelevant information (Francis & Love, 2020; Strauss & Francis, 2017). By using working memory processes, we are able to follow the conversation, actively participate, and take turns in the conversation (Edwards, 2007). However, working memory processes are limited and their greater use requires effort.

Although several behavioral and neural studies have attempted to explore the basis of listening difficulties, less is known about how background noise affects cognitive processes related to listening effort. Previous studies that presented participants with acoustically degraded speech (i.e., through filtering techniques) reported that sensory acoustic degradation represents an additional use of working memory processes that are likely to continue into memory storage phases and impair resources required for memory storage and language processes (Rabbitt, 2007; Piquado et al., 2010; Wingfield et al., 2016; Obleser et al., 2012). This is consistent with studies showing encoding words in the presence of background noise reduces the later recall performance of those words, even when the words were correctly heard (Kathleen Pichora-Fuller et al., 1995; Piquado et al., 2010). That is even though speech intelligibility is not impaired, the noise increases the load on working memory and reduces the capacity for encoding, storing, and further processing of words. These studies demonstrate that acoustic degradation and background noise affects higher-order cognitive processing that may lead to increased listening effort.

Various measures of listening effort have been proposed, such as behavioral, neuroimaging, and physiological measures. Assessment of listening effort has received much interest because it improves our understanding of daily-life difficulties in hearing-impaired individuals.

1.2 Measures of listening effort

1.2.1 Self-report

Self-report measures are subjective methods that rely on reporting how much effort the participants experience during a listening task (Johnson et al., 2015). The listener typically answers a question about how effortful the task was with a rating scale (Hart & Staveland, 1988). If they were paying adequate attention and motivated during listening, they would answer the questions based on all the energy and cognitive resources they spent listening and understanding the speech (Zekveld & Kramer, 2014). In other words, it could be argued that this method is a kind of holistic view of listening effort. Moreover, self-report measures are easy to use and quick (Bess et al., 2020).

Besides these advantages, there are some limitations to the self-reported measure of listening effort. Firstly, by this method, direct and continuous monitoring of listening effort during the task is not readily available because the current self-report methods can only assess listening effort after a speech comprehension task (Hornsby, 2013). Furthermore, a subjective measure of effort provides insight into a listener's experience of listening effort, but its scores may differ depending on the listener's judging standards of how difficult the task was. (McGarrigle et al., 2014).

1.2.2 Functional Magnetic Resonance Imaging (fMRI)

fMRI can be used to assess the changes in blood flow in the brain as a result of speech manipulations in effortful listening (Wild et al., 2012). A meta-analysis of auditory neuroimaging studies by (Alain et al., 2018) revealed a common pattern of activation in the frontal lobe in response to speech manipulation. They analyzed studies investigating speech perception in three independent types of speech manipulation paradigms: speech-in-noise (SIN); spectrally degraded speech using filtering techniques; and linguistic complexity (i.e., syntactic, and semantic intricacy). Based on SIN studies, greater effort is linked to the activation of brain regions including the left inferior frontal gyrus (IFG), left inferior parietal lobule, and right insula. Studies on spectrally degraded speech demonstrated increased activation in the insula on both sides of the brain and the left superior temporal gyrus (STG). Additionally, studies analyzing the complexity of language revealed activation in the left IFG, right middle frontal gyrus, left middle temporal gyrus, and bilateral STG. These findings suggest that there is a specialized region within the left IFG and that varied executive networks contribute to effortful listening (Alain et al., 2018).

The use of fMRI in listening effort research provides useful insights into the location of brain activity associated with a task. However, certain limitations exist. The acoustic noise generated by fMRI machines can impact the perception of auditory stimuli (Blackman et al., 2011). Additionally, the low temporal resolution of fMRI limits its ability to detect rapid task-evoked responses (Lystad et al., 2009).

1.2.3 Electroencephalography (EEG)

Exploring listening effort from brain activity recorded by electroencephalography (EEG) has received much interest (e.g., Dimitrijevic et al., 2019; Paul et al., 2021; Wisniewski et al., 2015). EEG is a non-invasive method that records cortical electrical activity through electrodes attached to the participant's scalp. EEG allows for real-time measurements of neural activity, enabling the assessment of early stages of information processing as well as later more cognitively controlled functions (Light et al., 2010). The frequency of the EEG signal (or electroencephalogram) ranges from 1 Hz to 200 Hz and has been divided into specific frequency ranges: delta (1-4 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (~ 33-200 Hz). Neural activity or oscillations belonging to each of these frequency ranges have been associated with different cognitive functions, such as working memory and attention. For instance, alpha oscillations have been related to inhibitory mechanisms, and theta oscillations to general cognitive load (Borghini et al., 2018; Khader et al., 2010). In the case of listening effort, changes in theta and alpha oscillations in speech-in-noise tasks may reflect changes in listening effort. For instance, a study led by (Dimitrijevic et al., 2019) showed that there is an association between alpha power (magnitude of oscillations) in the left inferior frontal gyrus and self-reported listening effort. The results confirm that increased alpha power is related to an increased effort rating, highlighting alpha power as a neural index of listening effort.

Estimating listening effort using EEG data might have some advantages over other methods such as fMRI. The EEG could be measured directly during the listening task with a high temporal resolution which allows studying of speech-brain association. When listening to speech, it has been shown that cortical activity continuously tracks the quasi-periodic features of speech, the phenomenon known as cortical speech tracking (Beier et al., 2021; Obleser & Kayser, 2019). Cortical tracking refers to the tendency of neural oscillations to align or phase-lock to speech features. These features of speech can be acoustic, such as the low-frequency amplitude modulation (envelope) of the speech signal associated with syllables (Peelle et al., 2013; Doelling et al., 2014), or linguistic representations generated in the listener's mind, such as syntactic phrase boundaries (Meyer et al., 2017; Ding et al., 2016). Cortical tracking plays a functional role in speech recognition.

A study by Dimitrijevic et al. (2019) found that speech-brain coherence in the 2-5 Hz range, which measures cortical speech tracking, is linked to the amount of effort required by participants to complete a digit-in-noise task. The experiment involved cochlear implant users who listened to a list of 9 digits while their EEG data was recorded. Participants rated their level of effort using the NASA Task Load Index (Hart & Staveland, 1988) on a scale of 1 to 10. The study concluded that low speech-brain coherence was linked to higher listening effort.

CHAPTER 2. AIMS AND THESIS OUTLINE

This thesis comprises three studies. The present chapter outlines the overall objectives of each individual study. Chapter 3 details the methods and materials that were used in each study. Chapter 4 summarizes the results of the studies. In Chapter 5, the outcomes of each study are discussed. Lastly, Chapter 6 offers a general conclusion to the thesis and highlights the key findings of the studies.

In this thesis, EEG was used to continuously record brain activity responses associated with listening to two types of speech stimuli that differ in their linguistic content: meaningful sentences and random word lists. The speech was presented in the background noise with different signal-to-noise ratios (SNRs). Participants listened to sentences and word lists and were asked to maintain them in memory and then respond. The EEG and behavioral data were recorded in two sessions. Study I used the data recorded in session one to study the effect of task manipulations on frontal theta oscillation during retention, which likely represents memory load. Study II aimed to study the reliability analysis of the neural and self-reported measures of listening effort with data recorded in both sessions. Study III aimed to study the effect of linguistic content and SNR on speech-brain association during the listening interval using the first session data set.

A details description of each study is presented in the following.

2.1 Study I

Speech in noise perception is a multi-stage process that involves a dynamic interaction of acoustic, linguistic, and cognitive factors (Shen et al., 2022). The present study aimed to investigate the interaction between linguistic violations (operationalized by sentences vs. word lists) and different levels of the acoustic challenge of noise (SNR of -9, -6, -3, and 0 dB) on listening effort during speech-in-noise perception.

As mentioned in the Introduction section, working memory mechanisms are the main contributor to effortful listening. Previous studies related to working memory tasks reported that maintenance of random word lists (presented in the quiet condition) elicited greater frontal theta ERS than normal sentences, indicating enhanced memory load regarding the maintenance of word lists (Meltzer et al., 2017; Bonhage et al., 2017). Other studies have shown that speech in noise perception increases theta band power in the frontal midline regions (Wisniewski et al. 2021; Wisniewski et al. 2015). Therefore in this study, we aimed to further study if frontal theta ERS is related to verbal working memory as a component of listening effort in speech-in-noise perception. if so, we expect that decreasing SNR to affect frontal theta ERS. We also expect the theta ERS to relate to subjective listening effort ratings.

Findings are reported in (Mohammadi et al., 2023a).

2.2 Study II

As reviewed in the introduction section, different subjective methods, e.g., self-reported, and objective methods such as EEG have been used to assess listening effort (McGarrigle et al., 2014; Peelle, 2018). In particular, changes in the amplitude of neuronal oscillations (alpha: 8-12 Hz and theta: 4-7 Hz) during the speech presentation and memory retention interval measured by EEG have been related to changes in listening effort (Ala et al., 2020; Paul et al., 2021; Dimitrijevic et al., 2019; Wisniewski et al., 2015). However, the reliability assessment of these measures is comparatively scarce. Reliability analysis examines the variability of a response over repeated measurements under constant experimental conditions (Downing, 2004).

The purpose of this study was to assess the test-retest reliability of self-reported listening effort and frontal midline theta (Fmθ) power in speech-in-noise recognition tasks. As increased Fmθ power during speech presentation and memory retention interval has been shown to reflect the increased listening effort, the reliability of the measure during these two intervals was assessed.

Findings are reported in (Mohammadi et al., 2023b).

2.3 Study III

Researchers have been studying cortical speech tracking to uncover the neural mechanism underlying speech perception. Cortical tracking of speech envelope mainly occurs in two frequency bands: theta (4-8 Hz) and delta (1-4 Hz). Theta band tracking is mostly linked to lower-level acoustic processing, such as syllable parsing, while delta band tracking pertains to higher-level linguistic information, such as words and word sequences. However, further research is needed to better understand the exact relationship between cortical tracking and acoustic and linguistic processing.

The impact of background noise on cortical speech tracking remains also a topic of debate among experts. While Ding & Simon (2013) reported that neural tracking of target speech was relatively unaffected by background noise, others have reported a significant decrease in target speech tracking as SNR decreases (Ghinst et al., 2019; Petersen et al., 2017). A study led by (Dimitrijevic et al., 2019) further highlighted this, showing that a decrease in speech-brain coherence or cortical speech tracking in a speech-in-noise task was associated with increased listening effort.

In this study, we further assessed the effect of linguistic information and background noise on speech tracking in speech-in-noise perception and its relationship with self-reported listening effort.

Findings are reported in (Mohammadi et al., 2023c).

CHAPTER 3. METHODS AND MATERIALS

This chapter gives an overview of the methods used in this thesis. In particular, it contains detailed information on participants, experimental design, data recording and analysis, and statistical analysis.

3.1 Experiment Setup

3.1.1 Participants

The experiment was conducted at Aalborg University in two sessions with identical conditions, separated by 6 ± 3 days. Thirthy-two healthy, normal-hearing, and native Danish speakers (13 females and 19 males, mean age = 24 ± 3 years). None of the participants reported a history of neurological or psychiatric illness. Written informed consent was obtained before participation and they received financial compensation. The study was conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of Northern Jutland, Denmark (N-20200061).

3.1.2 Speech Stimuli

The speech materials consisted of two types of speech: sentences with valid semantic and syntactic information and word lists with no syntactic and sentence-level semantic structure. Sentences, each made up of a subject, verb, numeral, adjective, and object, with the same syntactical structure but semantically unpredictable (e.g., in English: "Ulla owns five red flowers"). The mean duration of sentences was 2.22 ± 0.12 s. To create word lists, each sentence was split into different separate words. Each word list was created by a random combination of five separated words; for example (in English): "red four finds gifts two" and "six selected Henning bought had". The mean duration of word lists was 2.20 ± 0.16 s comparable to sentences. Final audio files of two speech types were masked by speech-shaped noise, by the modifying intensity of speech while keeping the intensity of noise constant, with SNRs of -9 dB, -6 dB, -3 dB, and 0 dB.

3.1.3 Stimulus Presentation

The experiment was presented in a factorial design with two independent variables: speech type (sentences and word lists), and SNR ($-9, -6, -3$, and 0 dB). A combination of two speech types with four levels of SNR gave a total of 8 conditions which were introduced in a block design (figure 1A). In each block, 25 trials were present (200 trials in total). Figure 1B shows a typical trial with four intervals starting with baseline continuing with listening and retention, and ending with response interval. In the response interval, participants were told to select the heard words verbatim in order. The intelligibility score was calculated as the percentage of correctly reported words in each trial. After completing each block, participants were asked to rate their level of listening effort on a scale of $1 - 10$ using the NASA Task Load Index (Hart & Staveland 1988) which is a self-reported measure. This approach was used widely in previous listening effort studies (Wisniewski et al. 2015; Paul et al. 2021; Dimitrijevic et al. 2019).

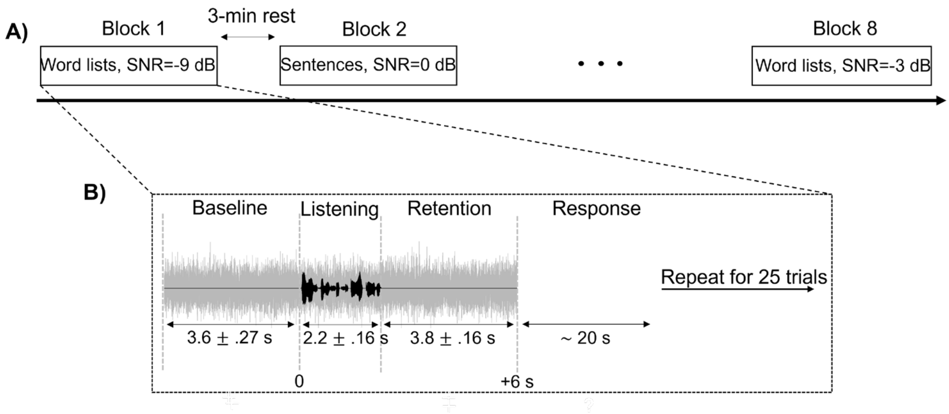

Figure 1. Overall experimental procedure. (A) Blocks were randomized in terms of speech type (word lists and sentences) and signal-to-noise ratio (SNR) $(-9, -6, -3$, and 0 dB), resulting in 8 conditions. After each block, participants had a 3-minute rest. (B) Depiction of a typical trial. (Mean $\pm$ SD)

3.2 EEG Recording and Preprocessing

EEG data were recorded by a g.HIamp biosignal amplifier (g.tec medical engineering GmbH, Austria) with 64 Ag/Ag-Cl channels and at a sampling rate of 1200 Hz. The left earlobe (A1) served as the reference point. To process the data, we utilized a custom MATLAB script and the EEGLAB toolbox (Delorme & Makeig 2004), filtering the EEG data from 0.5 to 40 Hz and downsampling it to 256 Hz. We then used the "Clean_rawdata" plugin in EEGLAB to detect and correct any bad portions of data that contained high-frequency activities Additionally, we used Independent Component Analysis (ICA) to remove any artifacts, including muscle and cardiac movements. The data was then re-referenced to the average reference, epoched to -3 to 6 seconds relative to the speech onset, and exported for analysis in BESA Research 7.1 (https://www.besa.de/). EEG datasets of one participant were excluded from the dataset due to an internal amplifier failure during recording.

3.3 EEG Data Analysis

Since each study used a different method to analyze the EEG data, we have presented the methods separately for each study below.

3.3.1 Study I

A complex demodulation method that was implemented in BESA Research was used to transform the epoched EEG data into a time-frequency representation. Spectral changes relative to baseline (-3 to 0) were quantified as a percent change. A power decrease relative to baseline indicates event-related desynchronization (ERD) and a power increase indicates event-related synchronization (ERS). Source localization was done by applying the multiple source beamformers (MSBF) method, as implemented in BESA Research, to time-frequency ranges of interest (retention interval: 3 - 6 s and theta frequency band: 4 - 7 Hz). we focused on retention interval to study memory load.

Statistical Analysis

Behavioral data (intelligibility and listening effort scores) were analyzed by a two-way repeated measure ANOVA (IBM SPSS Statistics 27 was used). Factors were speech type (sentences and word lists) and SNR (-9, -6, -3, and 0 dB). In the case that sphericity was violated, Greenhouse-Gierser correction was applied. For multiple testing problems, Bonferoni-Holm corrected p-values were reported.

The time-frequency-sensor data ($3-6$ s and $4-7$ Hz) were submitted to separate cluster-based permutation t-tests (paired, two-tailed, with 2000 permutations, cluster entry criteria (alpha) of 0.05) to test differences between sentences and word lists at each SNR. Additionally, cluster-based permutation one-way ANOVA tests (repeated measures, with 2000 permutations) were conducted on source-level data in sentences and word lists to test the effect of the SNR factor on theta ERS.

ROI Analysis

Since cluster-based permutation tests did not readily allow for testing interaction effect, region of interest (ROI) analysis was conducted to test the interaction effect of SNR and speech type. ROI was defined as a 7mm-radius sphere around the voxel of interest and averaged the values across voxels within the region (Meltzer et al. 2017). Then, on this data, a two-way repeated-measures ANOVA (in SPSS) was performed with the same factors.

Regression Analysis

To examine the relationship between effort rating and frontal theta ERS across various speech types, we utilized a linear mixed-effects model (LMM). The LMM incorporated SNR, recognition performance, and listening effort as fixed effects, and the participant as a random effect.

3.3.2 Study II

On the scalp time-frequency data, average spectral power across the 4–7 Hz frequency band and 3–6 s time window (retention interval) over ten frontal electrodes (AF3, AF4, F3, F1, FZ, F2, F4, FC1, FCZ, and FC2) were calculated as frontal midline theta (Fmθ) power.

Statistical Analysis

A two-way repeated measure ANOVA with the session (first and second), and SNR (levels: -9 dB, -6 dB, -3 dB, and 0 dB) as independent variables was applied to the intelligibility, listening effort, and Fmθ power values to test the main effect of a session and potential session×SNR interaction effect.

Reliability Analysis

Reliability analyses were conducted on the intelligibility, listening effort scores, and Fmθ power values for each condition (for two speech types at four SNR levels). Between-session relative reliability was assessed by an intraclass correlation coefficient (ICC) and absolute reliability was determined using the standard error of measurement (SEM) and Bland-Altman (BA) analysis (Bland & Altman, 1999). SEM was used to calculate the smallest detectable change (SDC), defined as the minimum amount of change in the score that can be interpreted as a real change for an individual rather than, potentially, the result of measurement error (Geerinck et al., 2019; Ries et al., 2009; Overend et al., 2010). A smaller SDC indicates a more sensitive measure.

3.3.3 Study III

Speech-brain phase locking value (PLV)

Speech audio signals were downsampled to 1200 Hz from 44.1 kHz before calculating the envelopes of each stimulus using MATLAB's Hilbert transformation. The resulting signal was downsampled to 256 Hz and then converted to the time-frequency domain through complex demodulation. To quantify cortical speech tracking, we used the phase-locking value (PLV) (Lachaux et al., 1999) in BESA Research to determine the consistency of phase difference between speech envelope and neural oscillations in various frequency bands. Additionally, we carried out PLV analysis on 100 random pairings of speech envelopes with EEG signal, which was averaged to generate random PLV as a baseline for each condition. For further analysis, the random PLVs were subtracted from the true PLV (i.e., correct speech-brain pairing).

Statistical Analysis

We conducted cluster-based permutation t-tests (paired, two-tailed, with 5000 permutations) to compare the PLVs of sentences and word lists across different SNR levels. The PLVs were averaged over the frequency and time interval of interest. We also performed separate cluster-based permutation ANOVA tests to examine the effect of SNR on speech-brain phase locking for both sentence and word list stimuli. We analyzed the potential interaction effect between SNR and speech type by averaging the PLVs of electrodes belonging to significant clusters of both SNR and speech type. Finally, on these data a 2 (speech types) × 4 (SNR levels) repeated-measure (RM) ANOVA was conducted, applying a false discovery rate (FDR) correction to address multiple comparison problems.

An LMM was used to examine the association between speech-brain PLV and listening effort for each speech type. SNR, speech intelligibility, and listening effort were the fixed effects in the LMM, while participants were the random effects.

CHAPTER 4. SUMMARY OF THE MAIN FINDINGS

This chapter presents a summary of the main results obtained for each study. In particular, study I contains behavioral (intelligibility and self-reported listening effort) and neural results regarding the effect of linguistic violation and SNR on frontal theta ERS. Study II consists of results of reliability analysis of intelligibility, effort rating, and Fmθ power. Study III presents the findings of speech-brian phase locking for sentences and word lists at different SNR levels.

4.1 Study I

Please refer to the original paper for detailed results (Mohammadi et al., 2023a).

4.1.1 Behavioral Performance

Figure 2 displays the average scores for intelligibility and listening effort. According to the statistical analysis for intelligibility scores (Figure 2A), there was a significant interaction effect ($p < 0.001$) between speech type and SNR, as well as significant main effects for both SNR ($p < 0.001$) and speech type ($p < 0.001$). Post-hoc pairwise comparisons revealed statistically significant differences ($p < 0.001$, Bonfferoni-Holm corrected), indicating that word lists had lower intelligibility scores than sentences, and increasing SNR improved intelligibility for both speech types. The observed interaction effect suggests that the degree of increase or decrease in intelligibility scores depends on the speech type. Specifically, increasing SNR from -9 dB to 0 dB raised the intelligibility of sentences to 98%, but only to 84% for word lists. This finding indicates that SNR has a varying impact on intelligibility in sentences and word lists.

Similar analyses were conducted for self-reported listening efforts (Figure 2B) that showed a significant interaction effect ($p < 0.001$) and a significant main effect of SNR ($p < 0.001$) and speech type ($p < 0.001$). The pairwise comparisons all showed statistical significance ($p < 0.001$, Bonferoni-Holm corrected), with higher listening effort for word lists than sentences. As SNR decreased listening effort scores increased in both speech types. The interaction effect showed that the effect of SNR on effort ratings differed across speech types. Particularly, decreasing SNR by 3 dB steps from 0 dB to -9 dB resulted in a more significant increase in effort rating at each step for sentences than word lists.

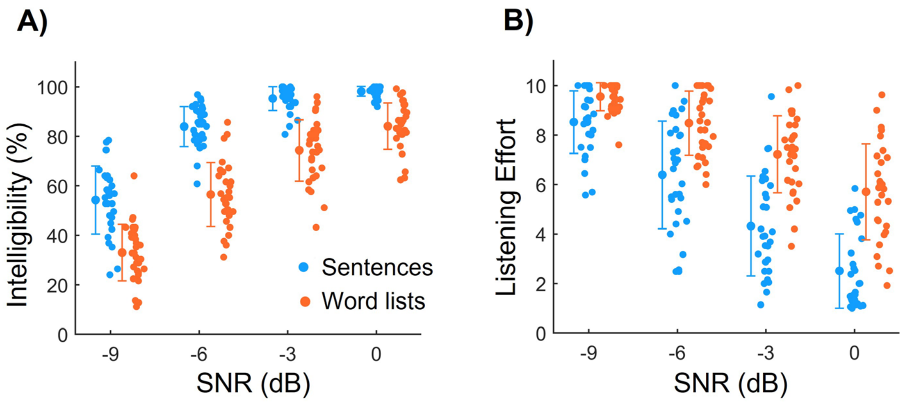

Figure 2. Behavioral responses. (**A**) Intelligibility is higher for sentences than for word lists and increases with increasing SNR. (**B**) Listening effort is higher for word lists than for sentences and decreases for both types of speech stimuli with higher SNR. Dots show the results from individual subjects, and the error bars show the standard deviation.

4.1.2 Neural Results

According to the results of the EEG source analysis, the experimental manipulations, which included varying background noise levels and presenting linguistic violations in the form of sentences and word lists, caused bilateral neural synchronization (ERS) in the theta frequency range (4-7 Hz) in the frontal cortex during the retention interval. The analysis revealed a significant difference cluster between word lists and sentences (p=0.032) located in the right middle frontal gyrus (right MFG) at Talairach coordinates x=31, y=46, z=23 when the background noise level was 0 dB (Figure 3B). During the memory retention interval, the ROI analysis (Figure 3C) showed an interaction between background noise and speech type on theta oscillations power in the right MFG (p=0.04). Post-hoc analysis in line with permutation testing revealed an increased frontal theta ERS for word lists compared to sentences only when speech was presented at the SNR of 0 dB (p=0.0012, Bonferroni-Holm corrected). No effect of SNR was observed. The LMM did not detect any significant relationship between frontal theta power and listening effort (p>0.05) in both sentences and word lists.

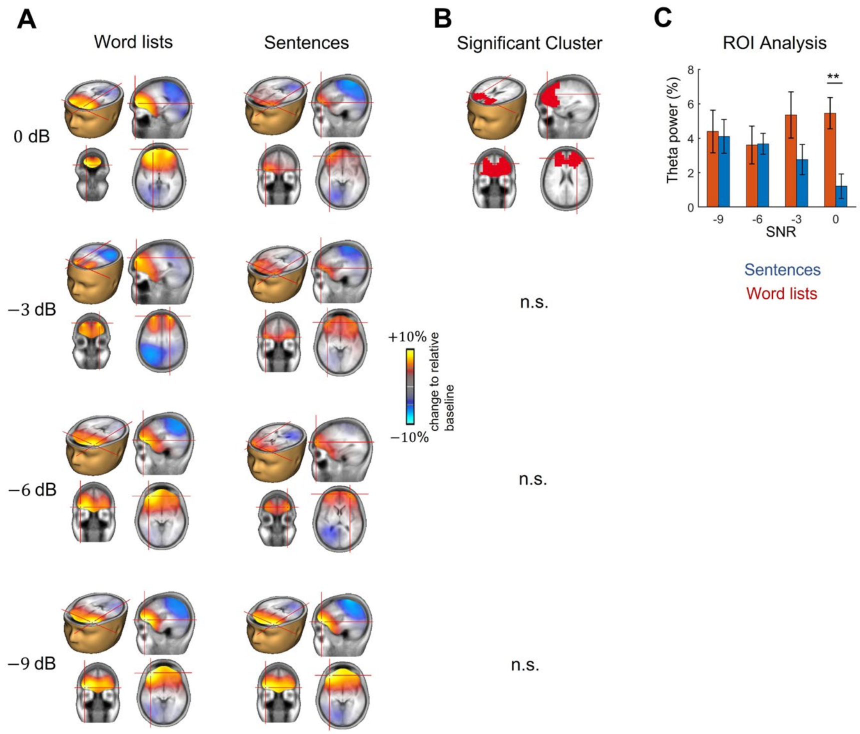

Figure 3. Source activity in the theta frequency band (4-7 Hz) during the memory retention interval. (A) Neuronal sources were localized in the frontal cortex in both word lists (left) and sentences (right) at different SNRs (0, -3, -6, and -9 dB). Crosshairs indicate peak source activations. (B) Cluster permutation test for significant differences between word lists and sentences in theta ERS. A significant difference was found at SNR 0 dB ($p = 0.032$), but not at other SNRs ($p > 0.05$). (C) Region of Interest (ROI) analysis. Mean theta ERS during retention in the right middle frontal gyrus (right MFG) ROI for word lists and sentences at different SNRs. The error bar represents the standard error of the mean. (**p=0.0012).

4.2 Study II

Please refer to the original paper for detailed statistical results (Mohammadi et al., 2023b).

4.2.1 Behavioral Data

The results indicate that there was a significant interaction effect between the session and SNR in sentence intelligibility scores (p=0.005). Further analysis revealed that there was a systematic bias (session 1 - session 2) at -9 dB ($p = 0.005$), -6 dB ($p < 0.001$), and -3 dB ($p = 0.015$), but not at 0 dB ($p = 0.90$). In word lists, there was a main effect of the session ($p < 0.001$), but no significant interaction effect was found between the session and SNR ($p = 0.27$). Additional post-hoc analysis showed that there was a difference between the sessions ($p < 0.001$).

The sentence condition showed a significant interaction effect between the session and SNR on listening effort ($p = 0.001$). Specifically, differences in effort ratings were observed between sessions at -6 dB ($p = 0.003$) and -3 dB ($p < 0.001$), but not at -9 dB ($p = 0.63$) or 0 dB ($p = 0.41$). In regards to word lists, a session × SNR interaction effect ($p = 0.027$) was also observed, with differences between sessions being found at -3 dB ($p = 0.003$), but not at -9 dB ($p = 0.13$), -6 dB ($p = 0.11$), or 0 dB ($p = 0.11$). Figure 4 and Table 1 provide relevant information such as BA plots, ICC, SEM, and SDC values for each condition along with their 95% CI.

4.2.2 Neural Data

Two-way repeated measures ANOVA performed on Fmθ power for word lists during the listening interval showed no relevant effects of the session, SNR, and session × SNR interaction effect. Analysis of sentences during the listening interval revealed a small main effect of SNR ($p = 0.04$) and no relevant effects of session and session × SNR interaction.

The ANOVA performed on Fmθ power for word lists during the retention interval showed no relevant main effect of the session, a main effect of SNR ($p = 0.03$), and no relevant session × SNR interaction effect. Concerning sentences, the analysis showed no relevant effects of the session, SNR, and session × SNR interaction. BA plots and ICC, SEM, and SDC values and their 95 % CI for each condition were listed in Figure 5 and Table 1 respectively.

	SNR – 9 dB		SNR -6 dB		SNR -3 dB		SNR 0 dB	
	Word lists	**Sentences**	**Word lists**	**Sentences**	**Word lists**	**Sentences**	**Word lists**	**Sentences**
Intelligibility Score (95% confidence interval)								
ICC	0.62 (0.19−0.82)	0.56 (0.2−0.77)	0.70 (0.30−0.87)	0.37 (-0.02−0.65)	0.62 (0.29−0.8)	0.21 (-0.09−0.5)	0.68 (0.44−0.83)	0.18 (-0.19−0.5)
SEM	5.6 (4.2−7)	8.2 (6.1−10.2)	5.6 (4.2−7)	5.2 (3.9−6.4)	6.7 (5.0−8.3)	3.1 (2.3−3.9)	5.2 (3.9−6.5)	2.4 (1.8−3)
SDC	15.5 (11.7−19.3)	22.7 (17.1−28.2)	15.6 (11.8−20)	14.3 (10.8−17.8)	18.5 (14−23)	8.7 (6.5−11)	14.4 (11−18)	6.7 (5.1−9)
Listening Effort Rating								
ICC	0.66 (0.41−0.82)	0.49 (0.18−0.71)	0.30 (-0.06−0.55)	0.26 (-0.04−0.54)	0.30 (-0.03−0.57)	0.42 (0.02−0.69)	0.36 (0.04−0.62)	0.50 (0.20−0.7)
SEM	0.37 (0.28−0.46)	0.98 (0.74−1.2)	1.28 (0.96−1.6)	1.76 (1.3−2.2)	1.39 (1.1−1.7)	1.26 (0.95−1.6)	1.51 (1.14−1.8)	0.96 (0.7−1.2)
SDC	1.02 (0.77−1.3)	2.7 (2−3.4)	3.5 (2.7−4.4)	4.9 (3.7−6.1)	3.8 (2.9−4.8)	3.5 (2.6−4.3)	4.2 (3.2−5.2)	2.7 (2−3.3)
Fmθ (Listening Interval)								
ICC	0.69 (0.45−0.84)	0.68 (0.43−0.83)	0.51 (0.19−0.74)	0.46 (0.12−0.70)	0.65 (0.38−0.81)	0.35 (-0.01−0.6)	0.62 (0.34−0.80)	0.16 (-0.2−0.5)
SEM	5.3 (4.0−6.7)	5.5 (4.1−7)	6.5 (4.9−8.2)	5.9 (4.4−7.4)	6.0 (4.4−7.4)	6.3 (4.7−8)	5.7 (4.2−7.1)	6.3 (4.74−8)
SDC	14.7 (11.0−18.5)	15.4 (11.5−19.3)	18.2 (13.6−22.8)	16.4 (12.3−20.6)	16.3 (12.2−20)	17.6 (13.2−22)	15.8 (11.9−20)	17.6 (13.1−22)
Fmθ (Retention Interval)								
ICC	0.62 (0.33−0.80)	049 (0.16−0.72)	0.39 (0.05−0.65)	0.42 (0.07−0.67)	0.33 (-0.01−0.61)	0.45 (0.10−0.70)	0.40 (0.10−0.69)	0.49 (0.15−0.7)
SEM	5.7 (4.3−7.2)	5.4 (4. 0 – 6.7)	6.5 (4.9−8.2)	4.3 (3.2−5.4)	7.9 (5.9−9.9)	5.0 (3.7−6.2)	7.1 (5.3−8.9)	4.5 (3.3−5.6)
SDC	15.9 (11.9−19.9)	14.9 (11.1−18.6)	18.2 (13.6−22.8)	11.9 (8.9−15.0)	21.8 (16.3−27.3)	13.8 (10.3−17.3)	19.8 (14.8−25)	12.4 (9.2−16)

Table 1. Intraclass correlation coefficient (ICC), Standard error of measurement (SEM), and smallest detectable change (SDC) for the intelligibility score, self-reported listening effort, frontal midline theta (Fmθ) power during listening and retention interval at different signal-to-noise ratios (SNRs).

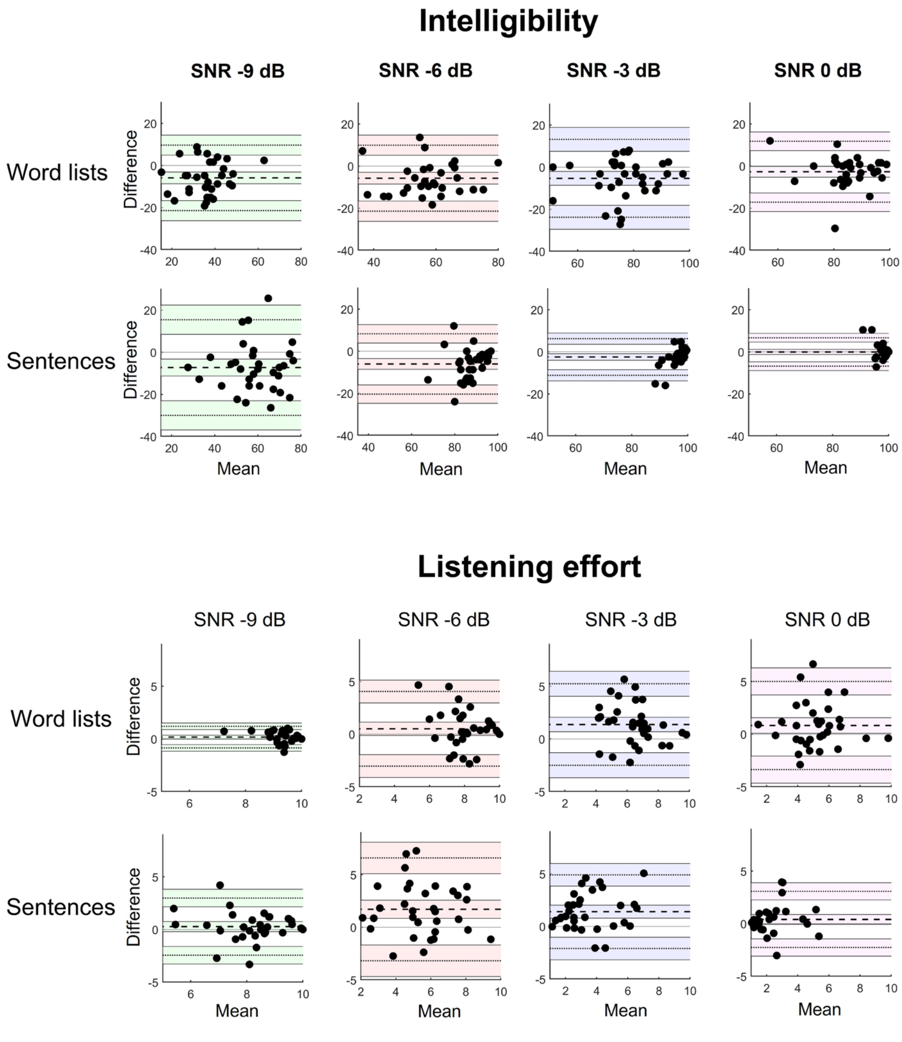

Figure 4. A) Bland–Altman plots of intelligibility and self-reported listening effort for word lists and sentences at a different SNR: -9 dB, -6 dB, -3 dB, and 0 dB. The dashed line indicates the bias between sessions, and the dotted lines are the limits of agreement (LoA), calculated as a bias ±1.96 times the standard deviation of the differences (SD_{diff}) in measurements between sessions. Shaded areas indicate the 95% confidence intervals of the bias and the LoA.

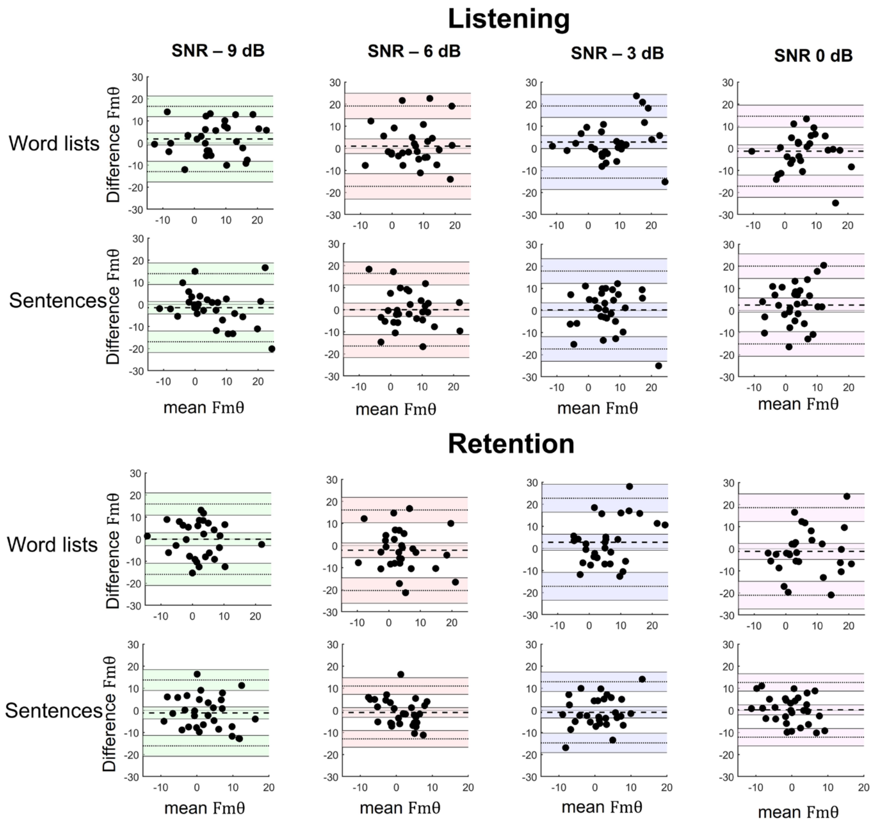

Figure 5. (A) Bland–Altman plots of $Fm\theta$ power at different signal-to-noise ratios (SNR)s for (top) listening and (bottom) retention intervals. The dashed line indicates the bias between sessions, and the dotted lines are the limits of agreement (LoA), calculated as a bias ±1.96 times the standard deviation of the differences (SD_{diff}) in measurements between sessions. Shaded areas indicate the 95% confidence intervals for the bias and the LoA. $Fm\theta$: frontal midline theta.

4.3 Study III

Please refer to the original paper for detailed statistical results (Mohammadi et al., 2023c).

We calculated the phase locking value (PLV) between brain activity and speech envelope in response to sentences and word lists at various frequencies (1-20 Hz) and time points (0-2 s). For each condition, we averaged the PLV values across the delta frequency band (1-4 Hz). Figure 6A displays two clear peaks for delta PLVs: the first peak from 0 to 500 ms and the second peak from 600 to 1100 ms. We averaged the PLVs (the difference between true and random values) within each interval (0-500 ms and 600-1100 ms) and then conducted cluster-based permutation t-tests to examine differences between sentences and word lists at each SNR. Across all SNRs, there were no significant differences between sentences and word lists in the first peak ($p > 0.05$). This suggests that this interval may be associated with an evoked response to speech onset that emerged for random PLV as well.

The results of the second peak showed a significant difference between sentences and word lists at SNR 0 dB on frontal and centro-parietal electrodes ($p = 0.028$ and $p = 0.035$, respectively) as shown in Figure 6C. However, no significant differences were found at other SNRs ($p > 0.05$). To evaluate the potential effect of SNR on the PLVs, separate cluster-based permutation ANOVA tests were conducted for both sentences and word lists. Results showed a significant main effect of SNR in sentences ($p = 0.024$) but not in word lists ($p > 0.05$). Further post-hoc analysis of the SNR effect revealed a significant difference in sentences between SNR of 0 dB and -9 dB ($p = 0.002$) as illustrated in Figure 6D. To test an interaction effect of speech type and SNR, we averaged the PLVs of those electrodes that belong to both clusters with significant differences between responses to sentences and word lists at 0 dB and between responses at 0 dB and -9 dB. Then averaged PLV was submitted to a two-way RM ANOVA. The result showed a significant speech type × SNR interaction effect ($p < 0.001$).

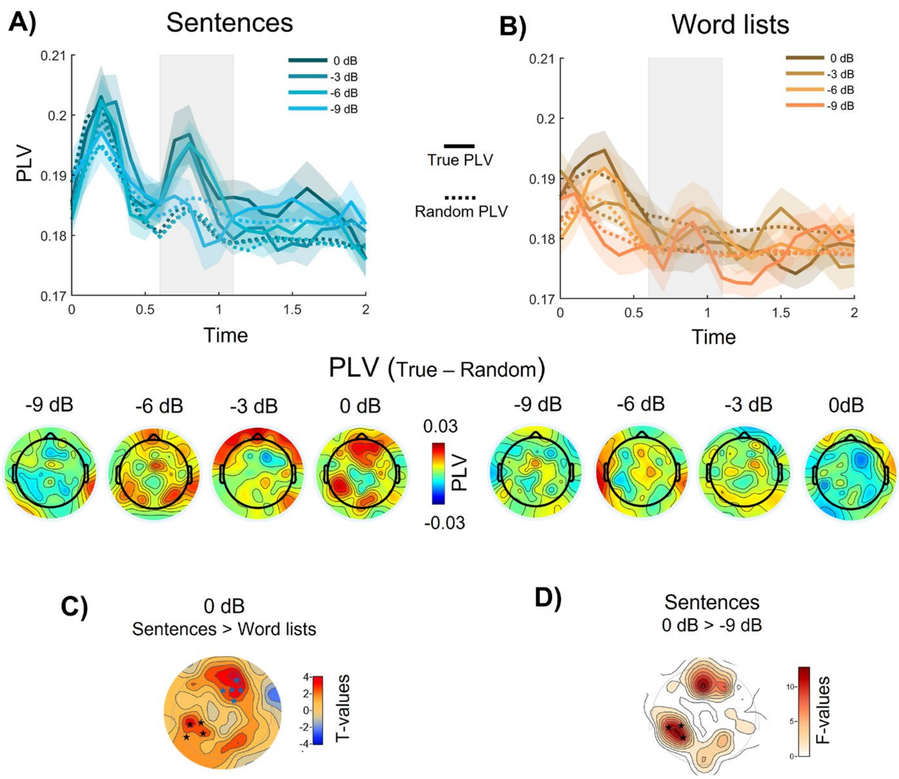

Figure 6. Delta-band speech-brain phase locking between speech envelope and EEG signals. (**A**) (top) The PLV values for correctly paired speech-brain signals (True PLV; solid line) and incorrectly paired signals (random PLV; dashed line) averaged across 62 electrodes and participants for different SNRs. The gray shaded area shows a time interval of interest (600-1100 ms). (bottom) Topographies of the delta PLVs averaged over the time interval of interest and all participants for different SNRs. (**B**) (top) PLVs for word lits averaged over electrodes and participants. (bottom) Topographies of PLVs averaged across the time interval of interest and participants. (**C**) Clusters of electrodes exhibited significantly higher PLVs in response to sentences as compared to word lists at the SNR of 0 dB. (**D**) A cluster of electrodes at the centro-parietal region showing a statistically significant increase of the PLVs in sentences at SNR of 0 dB than the SNR -9 dB (black stars; F is the F-value for the post-hoc test following the significant effect of SNR concerning PLVs for sentences).

The results of the LMM in sentences showed that listening effort had a marginally significant effect ($p = 0.06$) but SNR ($p = 0.13$) and intelligibility ($p = 0.98$) had no relevant effect. Similarly, there were no significant effects found in the word list, as shown in Figure 7.

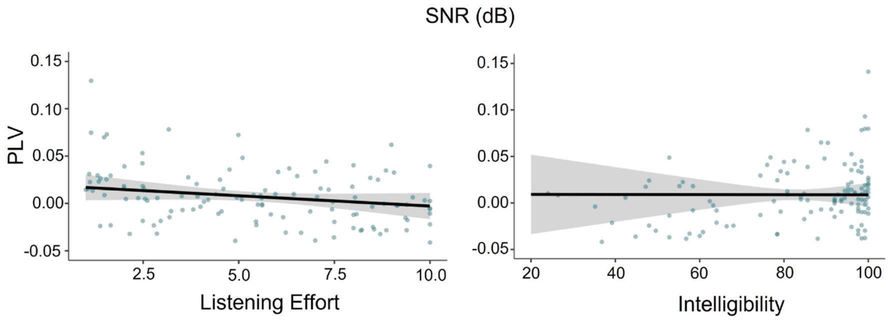

Figure 7. A linear mixed-effect model (LMM) was used to analyze PLVs in response to sentences. The results showed only a marginally significant effect on the listening effort (p = 0.06). The shaded areas in the graph represent the 95% confidence interval.

CHAPTER 5. DISCUSSION

This chapter provides a brief discussion of each study.

5.1 Study I

This study examined the effect of linguistic violations (operationalized as sentences vs. random word lists) and SNR and the interaction of these factors on frontal theta ERS during the retention interval. Dependent measures included intelligibility scores, a self-report measure of listening effort, and theta oscillatory power measured with EEG. The results revealed a significant interaction between linguistic violation and SNR on speech intelligibility scores, self-reported listening effort, and frontal theta ERS.

Behavioral results showed generally better word recognition performance for sentences than for word lists and the interaction effect showed that this increase in intelligibility score for sentences was more pronounced at lower SNRs. Sentences generally also showed lower listening effort ratings than word lists, and interactions indicated more reduction in effort for sentences at higher SNRs. It can be said that sentences are better recognized and require less effort than word lists because of their linguistic context (semantic and syntactic) (Mohammadi et al., 2023a). Indeed, sentential syntactic and semantic information allows words to be grouped into chunks that facilitate speech recognition and alleviate cognitive load As SNR decreased from 0 to -9 dB, intelligibility decreased for both sentences and word lists. The interaction showed that the decrease in intelligibility was pronounced in word lists as SNR decreased. For listening effort, also decreasing SNR by 3 dB per step from 0 to -9 dB enhanced the effort rating for both speech types. However, the interaction effect showed more increase in effort rating for sentences in each step. This suggests that SNR has a greater effect on sentences than word lists in effort ratings (Mohammadi et al., 2023a).

Results for frontal theta ERS during the retention interval indicated a difference between sentences and word lists (with more theta ERS for word lists than sentences) only at the easiest SNR (0 dB) and not at the other SNRs (-9, -6, -3 dB). As SNR decreased, the difference between word lists and sentences in theta ERS was no longer significant. Previous studies have shown that frontal theta ERS is substantially higher during the maintenance of word lists than sentences (Meltzer et al., 2017). This highlights that the greater memory load for word lists than sentences which is represented in increased frontal theta ERS might be related to the lack of linguistic information in word lists. The difference between word lists and sentences in theta power was not significant for other SNRs (-3, -6, and -9 dB), possibly because demand for sentences tends to increase with decreasing SNRs. However, statistical analysis on ROI data (theta ERS in rMFG) did not indicate significant changes for sentences between SNRs (Mohammadi et al., 2023a).

When it comes to the effort required for listening, the increase in frontal theta ERS is expected to coincide with the demand for verbal working memory. This means that when the SNR decreases, the frontal theta should also increase, as observed in the behavioral data that shows how SNR affected effort rating. However, our results showed that SNR did not have a significant effect on frontal theta ERS. It could be argued that theta ERS enhancement is generally related to demanding working memory conditions like image and digit maintenance, and it is not significantly affected by effortful listening in background noise. This is also supported by the lack of a significant relationship between theta ERS and self-reported listening effort (Mohammadi et al., 2023a).

5.2 Study II

The self-reported measurement of listening effort is based on a personal assessment of how challenging a task was, for instance in a speech-in-noise task (Johnson et al., 2015). As each individual has their own judgment standards, it is expected that the measured values may differ among participants. However, the crucial factor is how the scores vary between sessions or over time. If the effort score is consistent across sessions (Mohammadi et al., 2023b), it can be used in real-world scenarios to assess the level of effort required to perform the task (Cañete et al., 2022).

The reliability of self-reported listening effort was investigated in a study, which found acceptable variability between sessions. Using sentences at SNR 0 dB as an example, the results showed an SEM of 0.96 and an SDC of 2.7 points. The SEM value suggests that the difference between an individual's measured effort rating and the hypothetical true rating would be less than 1.88 (= 1.96*SEM) points for 95% of observations (Atkinson & Nevill, 2000; Bland & Altman, 1996), which is an acceptable difference given the 1-10 range for effort rating. Additionally, the SDC value indicates that an individual's effort rating would need to change by at least 2.7 points before it could be considered a true change, which is a small value and suggests that the self-reported measure is sensitive (Mohammadi et al., 2023b). These findings demonstrate the reliability of the self-reported measure for assessing listening effort.

For the neural measure of listening effort, our study aimed to evaluate the reliability of Fmθ power for both sentence and word lists, across various SNR levels and intervals (listening and retention). However, our analysis revealed a high level of variability in Fmθ power between sessions, across all conditions. For instance, taking the example of word lists at an SNR of 0 dB, the Fmθ power during listening was 7.23%, while during retention it was 6.26%. This suggests that subsequent participants tested in session one are likely to exhibit Fmθ power magnitudes of 7% during both listening and retention. However, based on our LoA analysis, there is a 95% chance that the retest Fmθ power will fall within a wide range of -10% to 21% for listening and -14% to 25% for retention, which may not be acceptable for practical use. This indicates that the use of Fmθ power as a reliable correlate of listening effort is questionable (Mohammadi et al., 2023b).

5.3 Study III

This study examined the effect of linguistic information and background noise on neural tracking of the speech envelope in a speech-in-noise task. The results showed an increase in delta-band PLV in response to sentences compared to word lists in the time interval of 600 – 1100 ms at SNR 0 dB. The increased PLV for sentences was significantly reduced as SNR decreased from 0 dB to -9 dB. The differences in delta PLV between sentences and word lists may be related to the top-down regulation of phase locking through high-level linguistic processing for sentences.

The sentences used in this study include a subject, verb, numeral, adjective, and object. It was observed that the phase locking between speech and brain activity was greater for sentences as compared to word lists, starting around 500 ms after speech onset. This timing is similar to the appearance of the second word (verb) in sentences, which is approximately 451 ms after onset (similarly, the second word in word lists appears around 443 ms after speech onset, but does not result in an increased phase locking). Hence, this increased phase locking at this particular latency for sentences may be connected to the cortical tracking of subject-verb structures (Ding et al., 2015). These findings suggest that the cortical activity is synchronized with internally constructed linguistic structures based on syntax (Peelle et al., 2013; Ding et al., 2015). Previous studies have also demonstrated the relation between cortical speech tracking in the delta frequency band and syntactic information (Meyer & Gumbert, 2018; Molinaro & Lizarazu, 2018). For example, Lu et al. (2022) and Coopmans et al. (2022) investigated the impact of sentential structure on neural tracking of word sequences and discovered that there was a significantly stronger delta-band neural response to regular sentences compared to random word lists, implying that delta-band neural responses are influenced by the compositional meaning of sentence structures (Mohammadi et al., 2023c).

Based on the LMM model results, the listening effort had the lowest p-value, suggesting that there might be an association between PLV and listening effort. However, this relationship was not statistically significant. Nonetheless, this trend aligns with the existing literature (Decruy et al., 2020; Dimitrijevic et al., 2019) which indicates that reduce in speech tracking might be related to increased listening effort.

.

CHAPTER 6. CONCLUSION

In this thesis, the effect of background noise and linguistic information on frontal theta oscillations and speech-brain phase locking was examined in speech-in-noise perception. The reliability of frontal theta power and self-reported listening effort was also assessed. When sentences and word lists were presented at high SNR (0 dB), increased theta ERS in word lists compared to sentences during retention interval was found. At this SNR level, an increased delta-band speech-brian phase locking was observed in sentences compared to word lists. These results highlight the role of linguistic information in sentence perception that showed a lower working memory load indexed by frontal theta power and high speech-brain phase locking in sentences. As SNR decreased from 0 dB to -9 dB, an increased trend in frontal theta ERS during retention and a significant decrease in delta PLV were observed in sentences and not in word lists, indicating the possible disruption effect of noise on linguistic information that enhanced perception load. Indeed, delta PLV in sentences showed a relationship with self-reported listening effort, that is increased effort was associated with decreased delta PLV. Reliability analysis showed low between-session variability in behavioral measures (intelligibility scores and self-reported listening effort), indicating a high level of reliability. Frontal midline theta power however showed high variability between sessions. Our study suggests that frontal theta power should be considered cautiously as a reliable neural correlate of listening effort, despite an acceptable effect size observed in distinguishing between sentences and word lists in retention interval.

CHAPTER 7. REFERENCES

Ala, T. S., Graversen, C., Wendt, D., Alickovic, E., Whitme, W. M., & Lunner, T. (2020). An exploratory Study of EEG Alpha Oscillation and Pupil Dilation in Hearing-Aid Users During Effortful listening to Continuous Speech. *PLOS ONE*, *15*(7), e0235782. https://doi.org/10.1371/JOURNAL.PONE.0235782

Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Human Brain Mapping*, *39*(7), 2695–2709. https://doi.org/10.1002/HBM.24031

Atkinson, G., & Nevill, A. (2000). Typical error versus limits of agreement. *Sports Medicine (Auckland, N.Z.)*, *30*(5), 375. https://doi.org/10.2165/00007256-200030050-00005

Beier, E. J., Chantavarin, S., Rehrig, G., Ferreira, F., & Miller, L. M. (2021). Cortical Tracking of Speech: Toward Collaboration between the Fields of Signal and Sentence Processing. *Journal of Cognitive Neuroscience*, *33*(4), 574–593. https://doi.org/10.1162/JOCN_A_01676

Bess, F. H., Davis, H., Camarata, S., & Hornsby, B. W. Y. (2020). Listening-Related Fatigue in Children With Unilateral Hearing Loss. *Language, Speech, and Hearing Services in Schools*, *51*(1), 84. https://doi.org/10.1044/2019_LSHSS-OCHL-19-0017

Blackman, G. A., & Hall, D. A. (2011). Reducing the effects of background noise during auditory functional magnetic resonance imaging of speech processing: qualitative and quantitative comparisons between two image acquisition schemes and noise cancellation. *Journal of Speech, Language, and Hearing Research : JSLHR*, *54*(2), 693–704. https://doi.org/10.1044/1092-4388(2010/10-0143)

Bland, J. M., & Altman, D. G. (1996). Statistics notes: Measurement error. *BMJ*, *312*(7047), 1654. https://doi.org/10.1136/BMJ.312.7047.1654

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160. https://doi.org/10.1177/096228029900800204

Borghini, G., Candini, M., Filannino, C., Hussain, M., Walsh, V., Romei, V., Zokaei, N., & Cappelletti, M. (2018). Alpha Oscillations Are Causally Linked to Inhibitory Abilities in Ageing. *The Journal of Neuroscience*, *38*(18), 4418. https://doi.org/10.1523/JNEUROSCI.1285-17.2018

Cañete, O. M., Nielsen, S. G., & Fuentes-López, E. (2022). Self-reported listening effort in adults with and without hearing loss: the Danish version of the Effort Assessment

Scale (D-EAS). *Disability and Rehabilitation*. https://doi.org/10.1080/09638288.2021.2022781

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414. https://doi.org/10.1016/J.TICS.2014.04.012

Coopmans, C. W., de Hoop, H., Hagoort, P., & Martin, A. E. (2022). Effects of Structure and Meaning on Cortical Tracking of Linguistic Units in Naturalistic Speech. *Neurobiology of Language*, *3*(3), 386–412. https://doi.org/10.1162/NOL_A_00070

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/J.COGPSYCH.2004.12.001

Decruy, L., Lesenfants, D., Vanthornhout, J., & Francart, T. (2020). Top-down modulation of neural envelope tracking: The interplay with behavioral, self-report and neural measures of listening effort. *European Journal of Neuroscience*, *52*(5), 3375–3393. https://doi.org/10.1111/EJN.14753

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. *Scientific Reports*, *9*(1), 11278. https://doi.org/10.1038/s41598-019-47643-1

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/NN.4186

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85 Pt 2*(0 2), 761–768.https://doi.org/10.1016/J.NEUROIMAGE.2013.06.035

Ding, N., & Simon, J. Z. (2013). Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *The Journal of Neuroscience*, *33*(13), 5728. 10.1523/JNEUROSCI.5297-12.2013

Downing, S. M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education*, *38*(9), 1006–1012. https://doi.org/10.1111/J.1365-2929.2004.01932.X

Edwards, B. (2007). The Future of Hearing Aid Technology. *Trends in Amplification*, *11*(1), 31–45. https://doi.org/10.1177/1084713806298004/ASSET/IMAGES/LARGE/10.1177_1084713806298004-FIG2.JPEG

Francis, A. L., & Love, J. (2020). Listening effort: Are we measuring cognition or affect, or both? *Wiley Interdisciplinary Reviews. Cognitive Science*, *11*(1). https://doi.org/10.1002/WCS.1514

Geerinck, A., Alekna, V., Beaudart, C., Bautmans, I., Cooper, C., de Souza Orlandi, F., Konstantynowicz, J., Montero-Errasquín, B., Topinková, E., Tsekoura, M., Reginster, J. Y., & Bruyère, O. (2019). Standard error of measurement and smallest detectable change of the Sarcopenia Quality of Life (SarQoL) questionnaire: An analysis of subjects from 9 validation studies. *PLoS ONE*, *14*(4). https://doi.org/10.1371/JOURNAL.PONE.0216065

Giuliani, N. P., Brown, C. J., & Wu, Y. H. (2021). Comparisons of the Sensitivity and Reliability of Multiple Measures of Listening Effort. *Ear and Hearing*, *42*(2), 465–474. https://doi.org/10.1097/AUD.0000000000000950

Ghinst, M. vander, Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Jousmäki, V., Hari, R., Goldman, S., & de Tiège, X. (2019). Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. *Journal of Neuroscience*, *39*(15), 2938–2950.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612. https://doi.org/10.1146/ANNUREV.PSYCH.49.1.585

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hornsby, B. W. Y. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, *34*(5), 523–534. https://doi.org/10.1097/AUD.0B013E31828003D8

Hsieh, L. T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *NeuroImage*, *85 Pt 2*(0 2), 721–729. https://doi.org/10.1016/J.NEUROIMAGE.2013.08.003

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *The European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/J.1460-9568.2002.01975.X

Johnson, J., Xu, J., Cox, R., & Pendergraf, P. (2015). A Comparison of Two Methods for Measuring Listening Effort As Part of an Audiologic Test Battery. *American Journal of Audiology*, *24*(3), 419. https://doi.org/10.1044/2015_AJA-14-0058

Kathleen Pichora-Fuller, M., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593–608. https://doi.org/10.1121/1.412282

Khader, P. H., Jost, K., Ranganath, C., & Rösler, F. (2010). Theta and Alpha oscillations during working-memory maintenance predict successful long-term memory encoding. *Neuroscience Letters*, *468*(3), 339. https://doi.org/10.1016/J.NEULET.2009.11.028

Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., Swerdlow, N. R., & Braff, D. L. (2010). Electroencephalography (EEG) and Event-Related Potentials (ERP's) with Human Participants. *Current Protocols in Neuroscience / Editorial Board, Jacqueline N. Crawley ... [et Al.]*, *CHAPTER*(SUPPL. 52), Unit. https://doi.org/10.1002/0471142301.NS0625S52

Lu, Y., Jin, P., Pan, X., & Ding, N. (2022). Delta-band neural activity primarily tracks sentences instead of semantic properties of words. *NeuroImage*, *251*, 118979. https://doi.org/10.1016/J.NEUROIMAGE.2022.118979

Lystad, R. P., Pollard, H., & Professor, A. (2009). Functional neuroimaging: a brief overview and feasibility for use in chiropractic research. *The Journal of the Canadian Chiropractic Association*, *53*(1), 59. /pmc/articles/PMC2652631/

Marsh, J. E., Ljung, R., Nöstl, A., Threadgold, E., & Campbell, T. A. (2015). Failing to get the gist of what's being said: Background noise impairs higher-order cognitive processing. *Frontiers in Psychology*, *6*(MAY), 548. https://doi.org/10.3389/FPSYG.2015.00548/ABSTRACT

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper.' *Http://Dx.Doi.Org/10.3109/14992027.2014.890296*, *53*(7), 433–445. https://doi.org/10.3109/14992027.2014.890296

Meyer, L., & Gumbert, M. (2018). Synchronization of Electrophysiological Responses with Speech Benefits Syntactic Information Processing. *Journal of Cognitive Neuroscience*, *30*(8), 1066–1074. https://doi.org/10.1162/JOCN_A_01236

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2017). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral Cortex (New York, N.Y. : 1991)*, *27*(9), 4293–4302. https://doi.org/10.1093/CERCOR/BHW228

Mohammadi, Y., Graversen, C., Østergaard, J., Manresa, J. B., Andersen, O. K. (2023a). Effects of background noise and linguistic violations on frontal theta oscillations during effortful listening. *Ear & Hearing,* Under Review.

Mohammadi, Y., Østergaard, J., Graversen, C., Andersen, O., & Manresa, J. B. (2023b). Validity and reliability of self-reported and neural measures of listening effort. https://doi.org/10.22541/AU.167850559.95803105

Mohammadi, Y., Graversen, C., østergaard, J., Andersen, O. K., & Reichenbach, T. (2023c). Phase-locking of Neural Activity to the Envelope of Speech in the Delta Frequency Band Reflects Differences between Word Lists and Sentences. *Journal of Cognitive Neuroscience*, 1–11. https://doi.org/10.1162/JOCN_A_02016

Molinaro, N., & Lizarazu, M. (2018). Delta (but not theta)-band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience*, *48*(7), 2642–2650. https://doi.org/10.1111/EJN.13811

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926. https://doi.org/10.1016/J.TICS.2019.08.004

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common α oscillatory network. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(36), 12376–12383. https://doi.org/10.1523/JNEUROSCI.4908-11.2012

Overend, T., Anderson, C., Sawant, A., Perryman, B., & Locking-Cusolito, H. (2010). Relative and absolute reliability of physical function measures in people with end-stage renal disease. *Physiotherapy Canada. Physiotherapie Canada*, *62*(2), 122–128. https://doi.org/10.3138/PHYSIO.62.2.122

Paul, B. T., Chen, J., Le, T., Lin, V., & Dimitrijevic, A. (2021). Cortical alpha oscillations in cochlear implant users reflect subjective listening effort during speech-in-noise perception. *PLOS ONE*, *16*(7), e0254162. https://doi.org/10.1371/JOURNAL.PONE.0254162

Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing*, *39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of Neurophysiology*, *117*(1), 18–27.

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex (New York, N.Y. : 1991)*, *23*(6), 1378–1387. https://doi.org/10.1093/CERCOR/BHS118

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37 Suppl 1*, 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

Piquado, T., Cousins, K. A. Q., Wingfield, A., & Miller, P. (2010). Effects of degraded sensory input on memory for speech: Behavioral data and a test of biologically constrained computational models. *Brain Research*, *1365*, 48. https://doi.org/10.1016/J.BRAINRES.2010.09.070

Rabbitt, P. M. A. (2007). Channel-capacity, intelligibility and immediate memory. *Http://Dx.Doi.Org/10.1080/14640746808400158*, *20*(3), 241–248. https://doi.org/10.1080/14640746808400158

Ries, J. D., Echternach, J. L., Nof, L., & Blodgett, M. G. (2009). Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Physical Therapy*, *89*(6), 569–579. https://doi.org/10.2522/PTJ.20080258

Shen, J., Fitzgerald, L. P., & Kulick, E. R. (2022). Interactions between acoustic challenges and processing depth in speech perception as measured by task-evoked pupil response. *Frontiers in Psychology*, *13*, 6099. https://doi.org/10.3389/FPSYG.2022.959638/BIBTEX

Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cognitive, Affective & Behavioral Neuroscience*, *17*(4), 809–825. https://doi.org/10.3758/S13415-017-0513-0

Wang, X., & Xu, L. (2021). Speech perception in noise: Masking and unmasking. *Journal of Otology*, *16*(2), 109–119. https://doi.org/10.1016/j.joto.2020.12.001

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(40), 14010–14021. https://doi.org/10.1523/JNEUROSCI.1528-12.2012

Wingfield, A., Tun, P. A., & McCoy, S. L. (2016). Hearing Loss in Older Adulthood. *Https://Doi.Org/10.1111/j.0963-7214.2005.00356.x*, *14*(3), 144–148. https://doi.org/10.1111/J.0963-7214.2005.00356.X

Wisniewski, M. G. (2017). Indices of Effortful Listening Can Be Mined from Existing Electroencephalographic Data. *Ear and Hearing*, *38*(1), e69–e73. https://doi.org/10.1097/AUD.0000000000000354

Wisniewski, M. G., Iyer, N., Thompson, E. R., & Simpson, B. D. (2018). Sustained frontal midline theta enhancements during effortful listening track working memory demands. *Hearing Research*, *358*, 37–41. https://doi.org/10.1016/J.HEARES.2017.11.009

Wisniewski, M. G., Thompson, E. R., Iyer, N., Estepp, J. R., Goder-Reiser, M. N., & Sullivan, S. C. (2015). Frontal midline θ power as an index of listening effort. *Neuroreport*, *26*(2), 94–99. https://doi.org/10.1097/WNR.0000000000000306

Wisniewski, M. G., Zakrzewski, A. C., Bell, D. R., & Wheeler, M. (2021). EEG power spectral dynamics associated with listening in adverse conditions. *Psychophysiology*, *58*(9). https://doi.org/10.1111/PSYP.13877

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*(3), 277–284. https://doi.org/10.1111/PSYP.12151

Zhang, L., Schlaghecken, F., Harte, J., & Roberts, K. L. (2021). The Influence of the Type of Background Noise on Perceptual Learning of Speech in Noise. *Frontiers in Neuroscience*, *15*, 484. https://doi.org/10.3389/FNINS.2021.646137/BIBTEX

ISSN (online): 2246-1302
ISBN (online): 978-87-7573-679-9
AALBORG UNIVERSITY PRESS