PERIODIC, APERIODIC, AND PHASE-LOCKED BRAIN ACTIVITY IN THE PERCEPTION OF CONTINUOUS SPEECH WITH DIFFERENT LEVELS OF DIFFICULTY AND INTELLIGIBILITY

Tuike Ekroos
Master's thesis
Department of Psychology
University of Jyväskylä
September 2023

UNIVERSITY OF JYVÄSKYLÄ

Department of Psychology

EKROOS, TUIKE: Periodic, aperiodic, and phase-locked brain activity in the perception of continuous speech with different levels of difficulty and intelligibility

Master's thesis, 52 p., 4 appendixes

Supervisor: Jan Kujala

Psychology

September 2023

ABSTRACT

Continuous speech is an important component of everyday communication between people. Previous research has suggested that both periodic and aperiodic brain activity occur during speech perception. In addition, brain activity has been suggested to be locked to the acoustic envelopes of speech signals. However, the neural responses during speech perception have so far been researched mainly using either individual speech stimuli such as syllables and words or sentences that do not fully correspond to those produced in natural speech situations. There has been little examination of continuous speech perception, and the different levels of difficulty and intelligibility typical of speech content have not been taken into account earlier.

The aim of this study is to examine the brain responses to speech when different levels of speech difficulty and intelligibility have been taken into account. This was done by studying the differences that can be found between different speech settings (difficult, easy, and vocoded speech) in different types of brain activity measures, i.e., periodic and aperiodic brain activity, as well as in the locking of brain activity to speech stimuli. The examination was done in the frontal, temporal and parietal cortical regions of the brain, that are known to be central to linguistic processing. The participants (N = 14) were healthy, native Finnish-speaking university students. In this study, speech intelligibility refers to the extent to which spectral and temporal detail of stimuli has been manipulated in vocoding while preserving the amplitude envelope whereas speech difficulty level refers to linguistic complexity of stimuli when the focus is on grammatical and lexical complexity.

The study reinforced the notion that speech perception is diverse in its neural basis. Statistically significant differences were found in all three types of brain activity measures when comparing the perception of speech stimuli with different levels of difficulty and intelligibility. However, the observed differences were partly contrary to expectations and to some extent internally contradictory and there was no evidence for any neural phenomenon alone to comprehensively explain the processing in the brain caused by speech stimuli of varying degrees of difficulty and intelligibility. The results can therefore be considered indicative of the complex nature of the neural processes involved in the perception of continuous speech and they underscore the need for additional work that considers larger sample size and more controlled stimuli.

Keywords: continuous speech, MEG, oscillations, arrhythmic brain activity, phase-locking

JYVÄSKYLÄN YLIOPISTO

Psykologian laitos

EKROOS, TUIKE: Periodic, aperiodic, and phase-locked brain activity in the perception of continuous speech with different levels of difficulty and intelligibility

Pro gradu -tutkielma, 52 s, 4 liitettä

Ohjaaja: Jan Kujala

Psykologia

Syyskuu 2023

TIIVISTELMÄ

Jatkuva puhe on tärkeä osa ihmisten välistä jokapäiväistä kommunikaatiota. Tähänastisessa tutkimuksessa on esitetty, että puheen havaintoon liittyy niin periodista ja aperiodista aivotoimintaa kuin aivotoiminnan lukkiutumista puheärsykkeisiin. Toistaiseksi puheen havaintoon liittyviä neuraalisia vasteita on tutkittu pääasiassa käyttämällä yksittäisiä puheärsykkeitä kuten tavuja ja sanoja tai lauseita, jotka eivät täysin vastaa luonnollisissa tilanteissa tuotettuja puheen osia. Jatkuvan puheen havaitsemista on tutkittu vähän, eikä puheelle tyypillisiä vaikeustasojen ja ymmärrettävyyden eroja ole otettu huomioon.

Tämän tutkimuksen tarkoituksena oli tarkastella jatkuvan puheen havaintoon liittyviä aivovasteita puheen eri vaikeustasot ja puheen ymmärrettävyys huomioiden. Tämä tehtiin tutkimalla eroja, joita on löydettävissä eri puhetilanteiden (vaikea, helppo ja vokoodattu puhe) välillä erilaisissa aivotoiminnan tyypeissä eli periodisessa ja aperiodisessa aivotoiminnassa sekä aivoaktivaation lukkiutumisessa puheärsykkeisiin. Tarkastelua tehtiin aivojen frontaalisilla, temporaalisilla ja parietaalisilla kortikaalisilla alueilla, joiden tiedetään olevan kielellisessä prosessoinnissa keskeisiä. Tutkittavat (N = 14) olivat terveitä, äidinkielenään suomea puhuvia yliopisto-opiskelijoita. Tässä tutkimuksessa puheen ymmärrettävyys liittyy siihen missä määrin puheärsykkeiden spektraalisia ja temporaalisia yksityiskohtia on manipuloitu vokoodauksessa kun taas puheen vaikeudella viitataan lingvistiseen monimutkaisuuteen päähuomion ollessa kieliopillisissa ja leksikaalisissa seikoissa.

Tutkimus vahvisti käsitystä, että puheen vastaanotto on neuraaliselta perustaltaan moninaista. Tilastollisesti merkitseviä eroja löydettiin kaikissa kolmessa tarkastellussa aivojen aktiivisuuden tyypissä verrattaessa vaikeustasoltaan ja ymmärrettävyydeltään erilaisia puheärsykkeitä toisiinsa. Löydetyt erot olivat kuitenkin jossain määrin odotusten vastaisia ja keskenään ristiriitaisia eikä näiden tulosten perusteella voida katsoa minkään tarkastelluista ilmiöistä selittävän eri vaikeustasoisen puheen tai ymmärrettävyydeltään eroavien puhetyyppien aivoissa aiheuttamia prosesseja aivoissa. Tuloksia voidaankin pitää suuntaa antavina antaen. Ne antavat viitteitä jatkuvan puheen havaintoon liittyvien aivotoimintojen moniulotteisesta luonteesta ja osoittavat tarvetta suurempaa otoskokoa ja kontrolloidumpia ärsykkeitä hyödyntäville jatkotutkimuksille.

Avainsanat: jatkuva puhe, MEG, oskillaatiot, rytmitön aivotoiminta, lukkiutuminen

CONTENTS

1. INTRODUCTION

1.1 Neural basis of auditory speech processing

The concept of 'speech perception' refers to the operations that transform auditory signals to linguistic representations that can interact with the mental lexicon (Holt & Lotto, 2010; Poeppel, 2015), while one form of it, auditory continuous speech, refers to the speech that we receive without visual cues (Madell et al., 2018). The auditory stimuli we receive vary a lot in terms of their different characteristics and complexity. Speech processing is an example of auditory perception, which involves dealing with complex sound stimuli. Since speech processing is an important part of human social interaction and has been found to be damaged by various brain-related diseases, developmental abnormalities, and brain injuries, understanding its underlying neural mechanisms is beneficial.

Functional localization mappings of speech and language in the cortex began in the mid-1800s when Broca and Wernicke introduced functional localization maps related to these functions. Since then, the views of the neural basis of language and speech as well as many other mental functions has shifted from localist to more dynamic, interconnective and integrative networks. In 1950's Penfied rejected the idea of static speech and language networks and instead proposed them to be based on dynamic connections, a notion that other authors confirmed by demonstrating the broad cortical distribution of language (Rahimpour et al., 2019). In 2007 Hickok and Poeppel presented a dual-stream model of speech processing arguing that a ventral stream processes speech signals for comprehension, while a dorsal stream maps acoustic speech signals to the articulatory networks in the frontal lobe (Hickok & Poeppel, 2007). This theory is still somewhat accepted.

As regards the neural basis of language and speech, one cannot bypass the discussion of the roles of the right and left hemispheres of the brain either. The topic has been discussed throughout the entire history of research of the neural basis of the language. As early as in 1836 Marc Dax presented this theory of localization of language in the left hemisphere, but this idea became better known later by Paul Broca (Manning & Thomas-Antérion, 2011). Since that the lateralization of neuronal processing related to various cognitive functions, including language and speech, has been extensively studied. Despite this, according to Assaneo et al. (2019) lateralization is still not fully understood. Even today, the hemispheres are considered to have different roles in speech processing, although there has been a shift from early theories emphasizing the left hemisphere in linguistic functions to more complex theories, in which different hemispheres of the brain are

considered to be involved complementarily in different sub-processes of language and speech, at least to some extent. An example of such a theory that emphasizes both hemispheres of the brain is The Asymmetric Sampling in Time theory (AST) presented by Poeppel (2003), according to which speech signals contain information at different time scales and speech perception takes place bilaterally in superior temporal cortex so that most of the language functions are lateralized to the dominant i.e., mostly left, hemisphere, but the analysis of speech is performed bilaterally. This theory has been considered to have received supporting evidence (Assaneo et al., 2019). Also, this study takes into account the possible differences between the hemispheres with respect to the studied phenomena.

In addition to examining the roles of the cerebral hemispheres, the role of different brain regions in the processing of speech perception has been studied extensively. Although localization theories emphasizing the importance of individual brain regions have largely been abandoned, it is still known that some brain regions are central to these processes. In particular, the temporal, parietal and frontal brain regions, and the connections between them can be considered interesting from the point of view of speech perception which is why this study has focused on looking at these areas.

Firstly, areas in temporal lobes have been the focus of attention from the beginning in the examination of neural processing of speech and even today these areas are considered central to the perception of speech. This is natural because the primary auditory areas are located bilaterally in them, and in addition the ventral pathway of the Dual stream model is largely located there. According to the Dual Stream model, the system diverges into two streams, ventral and dorsal from the early cortical stages of speech perception in the superior temporal gyrus (Saur et al., 2008). Ventral i.e. so called "what" stream is considered to transform acoustic sensory signals into linguistic representations while activity spreads from posterior to anterior temporal areas (Leonard & Chang, 2014). This pathway has been thought to be central for speech comprehension as it maps speech signals into conceptual and semantic representations (López-Barroso & de Diego-Balaguer, 2017). On the other hand, the so-called Wernicke area, which was originally localized to the superior temporal lobe used to be considered central to language understanding, but according to the prevailing view, the area is more involved in the production of speech (Binder, 2015). In their functional anatomical model of speech perception by Hickok and Poeppel (2000), the posterior half of the superior temporal lobe is seen as the primary site for constructing representations of auditory speech. Also, according to Schomers and Pulvermuller, especially in the left language dominant hemisphere, some of the activations are generated by the connections to the superior temporal primary auditory cortex in the afferent auditory hemisphere and spread from there to other areas that are specific to speech (Schomers & Pulvermüller, 2016).

Although often the left temporal regions have been considered more central to language processing than the right especially according to early authors, the right temporal regions have also been found to play a role in speech processing. According to Hickok and Poeppel (2000) also right superior temporal lobe participates in the perception of the speech and they claim that right and left hemispheres probably make different contributions to speech perception.

In addition to temporal areas, frontal and parietal brain regions are known to be activated in speech processing. The dorsal stream in the Dual Stream model projects from the superior temporal gyrus dorsoposteriorly to inferior parietal and posterior frontal lobe and is thought to participate in auditory-motor integration as it maps acoustic speech sounds to articulatory representations (Saur et al., 2008). The frontal lobe has traditionally been thought to be mainly related to the production of speech, but the motor areas there have been found to also participate in speech perception (D'Ausilio et al., 2012; Watkins et al., 2003; S. M. Wilson et al., 2004).

The development of the motor theories of speech perception began when Lieberman and Cooper presented their own theory in the 60s (Liberman et al., 1967). According to it, the perception of speech is closely related to the production of speech. In speech perception, differences have been observed in the activation of different motor brain regions, depending on whether the person sees the speaker's face or listens to the speech audio only. Audio-only speech listening has been found to activate the premotor but less the primary motor cortex whereas audiovisual speech perception has been found to activate both of these areas (Skipper et al., 2005). Glanz (Iljina) et al. (2018) used electrocorticography and noticed that the mouth motor region located in the superior ventral premotor cortex was activated during speech perception.

Although motor theories of speech perception no longer prevail in their original form, because, among other things, lesions in these areas have not been found to damage speech understanding, a new interest in the role of motor regions has arisen. Studies have found activation in the speech motor system areas during speech perception (Callan et al., 2010; Chevillet et al., 2013; Meister et al., 2007; S. M. Wilson et al., 2004). In addition, these areas may be of particular importance in situations where listening to speech is challenging. In their fMRI study Du et al. (2014) presented English phoneme tokens masked by noise and noticed activation in speech motor system areas including left premotor cortex and Broca's area, which, according to them, may indicate that discrimination in these areas may play a complementary role to sensorimotor integration.

The left inferior frontal gyrus, on the other hand, has been thought to play a role in the semantic processing of the language. In an experiment with a passive oddball paradigm in which spoken word pairs were presented, Shtyrov & Pulvermüller (2007) found that brain processes of semantic context integration can occur already at around 100 msec after the onset of spoken words

both in the superior temporal cortex and the left inferior frontal cortex. In addition, IFG has also been found to play a role in syntactic prediction of language (Matchin et al., 2017).

Evidence has also been obtained on the involvement of parietal brain regions and the somatosensory cortex in speech perception. While motor theories emphasized the importance of the activation of the motor brain regions in the speech perception, in their review Franken, Liu and Ostry point out that the connections between the somatosensory and auditory brain regions, as well as somatosensory feeds to the frontal motor brain regions, play an important role in the perception of speech (Franken et al., 2022). Thus, they propose a somatic role for the motor brain regions in the perception of the speech instead of a motor role. Parietal brain regions, like frontal and temporal ones, have been shown to be important for speech comprehension (Specht, 2014). Therefore, in this study, the examination of parietal brain regions has also been included when comparing speech of different levels of difficulty in the studied phenomena.

Although, among other things, the brain regions mentioned above have been shown to participate in speech perception, a complete definition of the functional neuroanatomy behind it can be considered difficult to establish. The functions of language are complex in themselves, since language processing including speech perception involves different subsystems involved in semantics, syntax, morphology, and phonology. In addition, speech is accompanied by non-verbal information, which the listener also processes. When listening to speech, we make use of various kinds of information, such as grammatical and semantic cues and prosody, including melody, rhythm and intensity (van der Burght et al., 2023). All these factors complicate the processing of speech. According to Hickok and Poeppel, reasons for the difficulty in revealing the neuroanatomy behind speech perception are that neural systems supporting these processes are different in different tasks and that systems that participate in the tasks presented under experimental conditions as for example syllable discrimination do not fully correspond to those that occur during natural speech perception (Hickok & Poeppel, 2000). In addition, according to them, the concept of speech perception can be understood in many ways and it can be tested behaviorally by various means.

Although more and more information has been gathered on the processing of language and speech and the brain regions involved since the days of Broca and Wernicke, more research is needed to understand the details of these complex processes. Connected speech is an important object of study when constructing this concept.

1.2 Connected speech

Connected speech or continuous speech refers to sets of sentences that are larger than individual syllables or words and that are often used in everyday communication between people. So far, neuroscience research has mainly focused on the processing of individual speech stimuli such as words and syllables and research on neural responses to continuous speech has been limited. According to Alexandrou et al. (2020) the research findings related to cortical entrainment and neural basis of the speech perception have been in general obtained using tightly controlled experimental paradigms using isolated sentences with constant speaking rate which does not correspond to natural speech as it occurs in real life. According to them, natural speech, compared to more controlled speech, has more disfluencies, less intelligibility, acoustic differences, quasi-rhythmicity, less salient temporal pattern, less constant intervals in the words of the sentences, less complete sentences, and less sharp acoustic edges.

However, there are several arguments for examining especially continuous speech. Firstly, continuous speech is a relevant part of verbal communication between people. Secondly, continuous speech is more than just the successive occurrence of several words in a row. In speech, it is known that there are changes in such a way that words are blended with each other. In speech, due to the motor behavior in speech production the elements are continuous while in written language they are discrete (Bäckström et al., 2022). So called speech segmentation problem arises as speech does not contain any acoustic analog of the blank spaces between words (Brent, 1999). The separation of words from each other is relevant both in the development of the child's language and in the learning of a foreign language. As a result, continuous speech cannot be viewed only as a combination of individual words. Thirdly, continuous speech contains additional information on how things are said by using tone of voice. Indeed, brain regions have been found that are involved in the perception of these language features. For example Sammler et al. (2015) revealed that prosody perception takes dual routes along dorsal and ventral pathways in the right hemisphere.

Due to the above considerations, when studying the neural bases of speech perception, the use of continuous speech as a signal, as we do in this study, provides additional information on the processing of speech perception in a way that the use of individual phonemes, syllables or words do not allow. Secondly, since connected speech varies in its difficulty, from simple words and sentence structures to more difficult ones in different contexts, taking this degree of difficulty and intelligibility into account, can provide additional information on how it affects the neural responses related to speech processing. So far, there has been little consideration of this dimension in previous neuroimaging studies. For these reasons, this study specifically examines connected speech and takes into account the different levels of difficulty and intelligibility of the language.

1.3 Periodic and aperiodic neural activity in the study of auditory speech processing

In addition to mapping the brain regions involved in speech perception and the connections between them, as well as the roles of different cerebral hemispheres in this perception, studies have examined various neural phenomena to understand their role in speech processing. This research dates back a long time. Already since Hebb's discoveries, it has been known that neurons do not work in isolation from each other but tend to act in synchrony (Hebb, 1949).

The electrical activity of the brain which can be represented in the form of brain waves, also known as neural oscillations, have been seen as rhythmic patterns characterized by peaks reflecting periodic fluctuations in post-synaptic potentials that rise above the broadband spectrum of neural signaling. The study of human brain oscillations began when Berger in 1920's measured electrical potentials between two electrodes placed at the scalp when the subject was asked to close her eyes and Berger observed around 10 Hz rhythmic activity later called alpha rhythm or Berger rhythm in the electrical potential (Jensen et al., 2014). Different oscillations were described already in 1920's and 1930's and were characterized by their dominant frequency (Hari & Puce, 2017). They can be divided according to frequency, or wavelength representing cycles per second (hertz, Hz) into five main categories, which are from the slowest to the fastest the delta, theta, alpha, beta and gamma rhythm. Since Berger's work attempts have been made to link these brain waves with different mental processes, as well sensory perception as higher-order cognition, including speech and language processing.

As a rule of thumb, the higher the frequency is, more alert and awake person is whereas delta brainwaves ($\sim$ 1–4 Hz) are typically considered to occur most during the sleep. Theta oscillations ($\sim$ 4-7 Hz), on the other hand, have been seen to be linked to tasks that require working memory or episodic memory (Sauseng et al., 2010; Tesche & Karhu, 2000). A burst of frontal theta activity is known to occur also in relation to processing information from an auditory stimulus (Picton & Picton, 2010). It has also been observed that cognitive load increases the power of theta oscillations (Snipes et al., 2022).

In addition, both theta and delta waves have been linked to speech processing. Various functional brain imaging studies have shown that auditory cortices track the time course of speaker's speech envelope at 0.5 Hz (delta frequencies) and 4–8 Hz (theta frequencies) while listening to speech (Bourguignon et al., 2012; Di Liberto et al., 2015; Ding & Simon, 2014). Theta and delta waves have been linked to prosodic and syllabic speech features and frequencies of both slow brain waves have been thought to track the speech rhythm. Furthermore, a link between delta and theta oscillations and speech intelligibility has been observed

with different roles assigned to theta and delta oscillations. Ding & Simon (2014) suggested that theta oscillations are related to encoding syllabic-level acoustic features critical for speech recognition while delta oscillations are more related to perceived acoustic rhythm rather than the phonemic information. They justify this, among other things, by the observation that speech modulations of 4-8 Hz have been found to be critical for intelligibility, while temporal modulations slower than this is frequency range important for music rhythm perception.

Giraud & Poeppel (2012), in turn, considered that the motor system and auditory system associated with human speech have developed over the course of evolution in such a way that they allow speech to be understood, and they also see neuronal oscillations as key phenomena for the auditory articulation of speech. Delta- and theta-band activations have been found to align to regulaties arising from speech characteristics like syllabic rates and phonemes and this alignment has been shown to be particularly strong in the auditory cortex (Keitel et al., 2018). Etard & Reichenbach (2019) presented continuous natural speech to the subjects both in their native language and a foreign language embedded in different levels of babble noise and noted that cortical speech tracking in the theta-band predicted speech clarity reflecting acoustics while tracking in the delta frequency predicted speech comprehension related to higher-level linguistic processing. Gross et al. (2013) used continuous speech both in intelligible and unintelligible, backward played form and noted stronger phase coupling between the speech envelope and brain oscillations in the intelligible compared than in the unintelligible condition in the left and right auditory cortex both in the delta- and theta-band frequencies. They also noted that gamma amplitude added complementary information to the theta phase suggesting that these mechanisms are partly independent of the other and can thus process complementary information. Moreover, the mutual information between neural activity in the delta band and the speech envelope has been suggested to be larger when subjects correctly understood speech in noise than when they misunderstood the target speech (Keitel et al., 2018).

Alpha oscillations ($\sim$ 8-14 Hz) have been seen as dominant oscillations in awake human brain (Halgren et al., 2019; Klimesch, 2012; Zhou et al., 2021) and they have been considered to be particularly associated with relaxed wakefulness with eyes closed and during REM sleep (Cantero et al., 2002; Clayton et al., 2018) reflecting inactive states (Lange et al., 2014). However, alpha oscillations have also been ascribed roles in various active processes such as visual (Busch et al., 2009; Samaha & Postle, 2015), auditory (Billig et al., 2019), multisensory perception (Ai & Ro, 2016; Bastiaansen et al., 2020; Cecere et al., 2015) as well as attention and inhibition (Arana et al., 2022; Klimesch et al., 2007). Generally, alpha oscillations have been studied in the visual and somatosensory systems where they are very prominent.

Although alpha waves are less noticeable in auditory modality and although these oscillations have been studied more in relation to for example visual functions, alpha waves have also been examined in relation to speech stimuli perception. Obleser & Weisz, (2012) presented individual words to the subjects and found in their EEG analyses that an increase in the spectral and envelope acoustic details of speech and speech understanding was associated with a decrease in the power of late (about 500ms post word onset) alpha waves at posterior channels. Ala et al., (2022) investigated alpha power as a measure of effortful listening in continuous speech by presenting babble noise with different signal-to-noise ratios. They found that alpha-band event-related synchronization (ERS) decreased with decreasing signal-to-noise ratio (i.e., increasing task demand) and that alpha-band ERS was positively correlated with task. According to them alpha power in continuous speech can correlate with how well speech is perceived.

Beta oscillations (13-25 Hz) have been investigated mostly in sensorimotor cortices and basal ganglia and have been seen to play a role especially in somatosensory processing and motor control (Barone & Rossiter, 2021; Khanna & Carmena, 2017) and also in working memory (Lundqvist et al., 2018). In addition, some connections between beta oscillations and speech processing have been suggested. Bidelman (2015) found correlation between induced activity in beta oscillations and strength of categorical speech perception, which refers to the ability to convert continuous variations in acoustics into discrete units, which is important in the perception of speech. Also Ou & Law (2019) examined roles of beta oscillations in speech perception. They suggested that these brain waves reflect individual differences in phonological representations as they found significant reduction in beta-band activity in a group of participants who were characterized as having non-distinctive perception compared to those who could perceive all tones but failed to distinguish between the two rising tones. In the first mentioned subjects, beta oscillations were also related to differences in discrimination sensitivity with lower beta-band activity associated with lower sensitivity.

Recently, even the fastest brain waves, gamma oscillations (in this study referring to oscillations > 25 Hz, in some studies >35 Hz) which can be recorded as well in many cortical as subcortical areas have aroused increasing interest among scientists (Başar et al., 2015). Like slower brain waves, gamma oscillations are believed to play roles in cognitive processes. The increase in the power of these rapid brain waves has been considered a key sign of information processing in the cortical regions (Muthukumaraswamy & Singh, 2013). Also links between gamma oscillations and several cognitive disorders like Alzheimer's disease and Fragile X syndrome (Mably & Colgin, 2018) as well as Parkinson's disease and schizophrenia (Guan et al., 2022) have been found. Visual stimuli have been found to induce gamma oscillations in visual cortex (Hermes et al., 2015). Furthermore,

transient bursts of gamma oscillations have been detected in response to auditory stimulation (Pantev et al., 1991) and this might been seen as the most widely studied human event-related synchronization in auditory system (Picton & Picton, 2010, p. 86).

As shown above, oscillations expressing periodic brain activity have been studied in numerous studies, and rhythmic neural actions behind them have been linked to many physiological and cognitive phenomena, including language and speech. However, the findings made so far have not been able to satisfactorily explain the complex processing, such as the perception of speech, and the connections found are mainly individual correlations. It should also be noted that in addition to this rhythmic activity, aperiodic, arrhythmic activity is noticeable in the brain. This scale-free, non-oscillatory network state is often called as 1/f. So far aperiodic activity has been studied less and has often been ignored or considered as an artefact. Because of this, the 1/f part of the brain activity is sometimes called noise. Almost until recent years, this type of brain activity was often excluded from the analyses (He, 2014). This can be considered problematic since connections have been found between aperiodic components and various phenomena. According to the results of the resting-state MEG study by Ibarra Chaoul & Siegel (2021), aperiodic neural activity is correlated in the entire cortex region and, in addition, the structure of this correlation is similar but distinct from the structure of oscillatory neural correlation. Based on this, an assumption can be made, according to them, that aperiodic activity acts as a robust marker of interactions in cortical networks, and that the examination of both periodic and aperiodic signal components can provide information on neural correlations. The aperiodic component also reflects physiological information (Donoghue et al., 2020) and it has been associated with for example aging (Voytek et al., 2015), neuropsychiatric diseases like schizophrenia (Peterson et al., 2018), ADHD (Ostlund et al., 2021), and cognitive functions like memory consolidation (Helfrich et al., 2021).

While periodic components appear as spectral peaks and have been considered as oscillations, an aperiodic part usually follows a power law $P \propto 1/f^{\beta}$ (Miller et al., 2009) also called 1/f where the power at the signal frequency decreases as a function of the frequency (f) (Voytek et al., 2015). Thus, low frequencies have large amplitudes and high frequencies have low amplitudes (Gilden, 2001). While the measures used in the study of periodic brain activity are frequencies, powers, and bandwidths, 1/f can be parametrized with the exponent and offset. Offset parameter reflects uniform shift of power across frequencies (Donoghue et al., 2020). It is probably not without meaning as increased offset has been shown to correlate with increased neuronal population spiking activity (Miller et al., 2009, 2014; Voytek & Knight, 2015). In addition, the exponent reflecting the pattern of aperiodic power across frequencies (Donoghue et al., 2020) has shown to be related to for

example excitation-inhibition balance (Gao et al., 2017) and diseases like schizophrenia (Molina et al., 2020), epilepsy (van Heumen et al., 2021) ja ADHD (Ostlund et al., 2021).

Indications have been found that 1/f like brain activity is not without functional meaning. For example, it has been observed that the presentation of sensory stimuli modulates the power-law exponent (Boustani et al., 2009; He et al., 2010). Associations between aperiodic brain and predictive processing about upcoming stimuli have also been claimed to exist (Maniscalco et al., 2018). Recent advances in parameterizing neural power spectra and the development of methods such as Spectral Parameterization Resolved in Time (SPRiNT) (L. E. Wilson et al., 2022), FOOOF (Fitting Oscillations & One-Over-F) and IRASA (Irregular Resampling Auto-Spectral Analysis) (Gerster et al., 2022) have made it possible to consider the periodic and aperiodic signals separately. For this reason, the FOOOF method was used in this study.

Although the research and methods of aperiodic brain activity that make this research possible have advanced, when it comes to language processing and perception of speech discussed here, the research of aperiodic brain activity is still in its infancy. However, some indicative results have already been obtained which underscore the need for additional work. Cross et al. (2021) suggested that $1/f$ activity plays an important role in higher order cognition like language processing, as in their artificial grammar learning task, learning-related aperiodic activity interacted with oscillatory power to modulate behavioral performance and the 1/f slope became steeper during the learning of complex rules. Another language-related finding is done by Dave et al. (2018) who found that 1/f activity was a predictor of N400 effects of successful lexical prediction while participants read two-sentence discourse passages. In terms of speech perception, the role of aperiodic brain activation remains unknown. However, findings such as those mentioned above suggest that aperiodic brain activation should not be considered just as a meaningless noise that is removed from the analysis. Instead, it can be fruitful to analyze its role together with oscillations in relation to different cognitive processes, including speech perception. Because of this, the present study does not limit itself to examining only rhythmic brain activity, but also examines rhythmless brain activity in the form of exponent and offset values.

1.4 Temporal structures of the speech and brain activity's entrainment to it

The oscillations have been considered to express the rhythm of brain activity due to the activity of neurons. Rhythmicity has also been presented to be typical of speech. The regularity of the rhythm of speech is suggested to play a key role in speech comprehension, which is why the envelope of the speech stimulus has been examined largely from the perspective of speech perception (Poeppel et al.,

2020). By using modulation spectrum analysis Ding et al. (2017) found that there is highly consistent speech modulation spectrum in different languages and speech produced in different manners and that speech shows peaking at around 5 Hz. In line with this finding Varnet et al. (2017) used different analyses to describe temporal structures of the speech signal (the amplitude-modulation and frequency-modulation spectra) and found similar results. According to Poeppel and Assaneo speech has a regular rate when it goes to the ear and this regular envelope of speech sound is produced by the natural biomechanical frequency of jaw movements (Poeppel et al., 2020). Several other researchers have suggested that for speech, temporal modulations resonate at around 4 to 5 Hz.

However, although the idea of rhythmic speech seems to prevail, there is no watertight consensus related to the temporal properties of speech. The presence of rhythmicity of the speech have even been denied (Nolan & Jeon, 2014). On the other hand, it has been suggested that the rhythmicity of speech depends on the speech stimuli used. According to Alexandrou et al. (2020) clearly rhythmic structure is typical to isolated sentences and read-aloud texts, not for natural connected speech. They have shown that the power spectrum of spontaneously produced speech is quite flat and is characterized by a 1/f trend (Alexandrou et al., 2016). Findings regarding 1/f fluctuations in speech is not new as Voss and Clarke discovered it back in the 1970s. (Voss & Clarke, 1975). Thus, the study of the neural basis of speech perception should consider the features of the stimuli used. If the neural basis for the perception of natural connected speech is to be investigated, the stimuli presented must be as close as possible to the speech produced in real-life speech situations, not individual words, or phrases.

It has been suggested that there is a connection between the rhythmic activity in the speech and in the brain. It has been observed that auditory cortical activity is entrained – meaning tracking or locking on – to the envelope of speech signal as subjects listen to speech. This entrainment has been observed in slow oscillations in both the delta and theta bands (Ding & Simon, 2012; Luo & Poeppel, 2007; Zion Golumbic et al., 2013). This neural envelope tracking in the auditory cortex has been suggested to have a connection to speech comprehension (Doelling et al., 2014; Luo & Poeppel, 2007; Peelle et al., 2013). Locking of brain activity to speech has been examined by comparing the locking caused by different speech stimuli with each other by varying the intelligibility of speech. Ding et al. (2014) found that cortical entrainment to natural speech is robust to background noise whereas entrainment to vocoded speech is not. According to them the spectral-temporal fine structure is degraded in noise-vocoded speech and is critical to segregating speech from noise. However, the connection between cortical entrainment and speech comprehension has also been met with reservations. Alexandrou et al. (2020) call for a critical view of the claim that the cortical entrainment is a prerequisite for speech perception and comprehension and that studies should use

natural, near real-life speech instead of for example isolated sentences or read-aloud texts featuring clearer rhythmic patterning to consider the role of the cortical entrainment. In their opinion it has not been verified experimentally that cortical entrainment really takes place during speech perception and according to them to call findings as cortical entrainment the coupling between ongoing cortical oscillatory activity and the acoustic amplitude envelope should persist over time. They suggest that the phenomena often referred to as cortical entrainment is actually a repetition of evoked potentials that result from by the acoustic edges in the speech (Alexandrou et al., 2020).

The locking of brain activity in the speech signal still requires further research. Therefore, in this study, in addition to rhythmic and arrhythmic brain activity, the third phenomenon to be studied is this locking of brain activity in speech. In accordance with the requirements of Alexandrou et al. (2020), a near real-life speech stimuli are used in the study.

1.5 Investigation of brain processes of speech perception using brain imaging methods

The development and improved availability of brain imaging techniques over the past decades has brought new opportunities when studying the neural responses of various functions. Functional imaging methods can be used to study in vivo the neural events underlying human sensorimotor and cognitive functions. These methods can be divided into two categories: electrical and magnetic methods and hemodynamic and metabolic methods. Electroencephalography (EEG) and magnetoencephalography (MEG) are examples of the former, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) of the latter. fMRI and PET have high spatial resolution but poor temporal resolution whereas EEG and MEG have high temporal but lower spatial resolution (Gosseries et al., 2008). Temporal resolution in MEG is in millisecond scale (Hari, 2011) whereas in fMRI it is from 1 to 4 seconds and in PET 30 seconds (Varvatsoulias, 2013).

MEG records magnetic fields generated by electric currents in the brain whereas EEG measures brain electrical fields. Because ionic currents are the main sources of both signals, EEG and MEG are quite close methodologies. However, MEG has some advantages over EEG. Because of less spatial smearing it has better spatial resolution when it comes to separating cortical sources, it is selective to fissural cortical activity (Hari, 2011), and it has been found to be more sensitive to the majority of cortical sources compared to EEG (Piastra et al., 2021) whereas some very deep sources are more reliably picked up by EEG than MEG (Hari, 2011).

According to Destoky et al. (2019) MEG is better than EEG in connected speech listening research as recordings in latter need to be around 3 times longer than in MEG to uncover significant speech tracking. Moreover, in MEG both signal-to-noise and spatial resolution are higher than in EEG (Fred et al., 2022). Nevertheless, one of the biggest limitations of MEG can be considered

the presence of physiological and technical disturbances like cardiac signals and eye movements (Mantini et al., 2008). These artifacts can be 10-20 times larger than evoked signals (Kawakatsu, 2003). However, in recent years, numerous methods have been developed that can be used to remove these artifacts.

Underlying neural mechanisms in cognitive functions happen in short time scales. For example, speech perception happens in some hundreds of milliseconds. Therefore, MEG and EEG have been thought to be the only noninvasive methods that are sufficient to estimate speech brain tracking (Destoky et al., 2019). In speech perception studies, EEG has been used significantly more than MEG due to its better availability. The availability of MEG hardware is limited by its high price. However, all the above methods have been used in the review of speech perception.

In this study MEG as a time-sensitive method was used. The choice was supported by the fact that the study looked at cortical phenomena, not deep sources. Methods to remove artifacts, Temporal Signal Space Separation, Oversampled Temporal Projection, and Independent Component Analysis, were used to reduce noise caused by the method.

1.6 Research questions

Numerous studies have sought to clarify the neural basis of speech perception. This investigation has been done specifically by looking at rhythmic brain activity. In the case of different oscillations, some links have been found to the processing of speech perception, but these connections cannot be considered to form a comprehensive conception explaining these activities. As for arrhythmic brain activity, the examination of the neural basis of speech has been very limited.

The research carried out so far has been overshadowed by the nature of the stimulus used. The use of natural speech corresponding to speech produced in natural speech situations has been limited in studies. Most of the studies have used individual syllables, words, or phrases. Using such stimuli cannot be able to reliably investigate the brain basis of connected speech. In addition, the differences in the level of difficulty and intelligibility typical of natural speech have so far been largely ignored. Therefore, this study wanted to expand on previous research by using natural continuous speech as stimuli. The aim was to look for possible differences in three brain activation phenomena, periodic and aperiodic brain activation, and the locking of brain activation in a speech signal when presenting speech stimuli of varying degrees of difficulty and intelligibility. The difficulty and intelligibility of speech were taken into account by comparing both easy and difficult as well as easy and vocoded speech stimuli.

In this study, speech intelligibility refers to the extent to which spectral and temporal detail of stimuli has been manipulated while preserving the amplitude envelope whereas speech difficulty level refers to linguistic complexity of stimuli when the focus is on grammatical and lexical complexity.

There were three research questions in this study:

Research question 1. What are the differences between perceiving different types of continuous speech (easy vs. difficult and easy vs. vocoded speech) in periodic brain activity?

Research question 2. What are the differences between perceiving different types if continuous speech (easy vs. difficult and easy vs. vocoded speech) in aperiodic brain activity?

Research question 3. What are the differences between different types of continuous speech (easy vs. difficult and easy vs. vocoded speech) in the locking of brain activity into speech stimuli?

As previous research on continuous speech has been limited with regard to the above-mentioned phenomena, no actual hypotheses were presented related to the research questions. However, cautious assumptions could be made that easy speech might differ from difficult (speech difficulty) and easy speech from vocoded (speech intelligibility), generating modulations in different neural mechanisms in different areas of the brain.

The phenomena of brain activity in question were considered in the frontal, parietal and temporal regions of the brain since these areas are known to be central to speech perception. The aim was to find out what kind of differences were observed in brain activation when comparing different levels of speech difficulty (easy vs. difficult) and, on the other hand, when making a corresponding comparison between different levels of intelligibility (easy vs. vocoded). It was hoped that this study would provide additional information on the features of brain activity associated with the perception of natural speech content with different levels of difficulty and intelligibility.

2. METHODS

2.1 Participants

The participants (N = 14) were healthy, native Finnish-speaking university students (10 women and 4 men; mean-age = 23,3-year, range = 20-33 years) with normal or corrected-to-normal vision and hearing, not diagnosed with neurological diseases, epilepsy, depression, or anxiety disorder. One of the participants was left-handed, 13 were right-handed. All participants gave their informed written consent before taking part in the experiment. Approval for the project was given by the Human Sciences Ethics Committee of the University of Jyväskylä, Finland.

2.2 Procedure

Before the experiment, five head position indicator coils were attached to each participant's scalp to determine their head position. Polhemus Fastrak 3D Digitizer and Quad Sensor Motion Tracker - device (Polhemus, Vermont, USA) was used to determine the coil locations. Right and left preauricular points and nasion point were used as anatomical landmarks. In addition, blinks and eye movements during the MEG experiment were measured using electro-oculography (EOG).

During the measurements, the participants were seated with the head covered by the MEG helmet. They were asked to minimize their movements in order to reduce the artifacts in the MEG measurement. A screen was placed in front of the subjects at one meter from their faces, on which, using the Presentation software, a fixation cross and questions with answer options were projected. The subject was given earbuds in each ear to present the stimuli.

Measurements were done in a magnetically shielded room at the Centre for Interdisciplinary Brain Research (CIBR), Jyväskylä University. The brain activity related to perception of continuous speech was recorded using an Elekta Neuromag Triux system (Elekta Neuromag, Finland), which is a 306 channel MEG system. It combines both planar gradiometers and magnetometers. In this research only planer gradiometers were included in data analysis because they are less sensitive to interference located distantly from the sensors (Mohseni et al., 2011), as they detect less environmental noises than magnetometers, and as they are not sensitive to the uniform fields (Sun et al., 2018).

Before the actual experiment, the subjects were presented with a familiarization scenario. This scenario was also used to adjust the volume. The actual experimental setup consisted of three conditions of speech perception (easy, difficult, and vocoded speech). Time of the total experiment was about half an hour. All subjects were presented with the same three sets of stimuli, easy, difficult, and vocoded continuous speech. The order of presentation of the conditions was randomized for each subject.

The setup included the performance of 15 recordings, with each condition consisting of the presentations of five samples of continuous speech. The average length of the samples was 79 s (range = 62 - 103 seconds). The samples consisted of Finnish-language speech by three different men and two women. The difficult condition consisted of 5 audio clips from radio programs made for adults and easy condition of 5 audio clips from radio programs made for children. The audio clips were obtained from Yle Audio (Appendix A) and were in Finnish. The topics of the programs were related to science and nature.

In the easy and difficult condition, the presentation of each audio file was followed by two multiple-choice questions about the content of clips. In the vocoded condition the subjects were asked to recognize which target word was intelligible. The subjects answered the multiple-choice questions using a four-button device compatible with the MEG device. Asking questions was used to keep participants in an alert listening mode during the presentation of stimulus sequences.

There were pauses between the presentations of the different conditions. The subject was asked about his willingness to take a break or immediately proceed to the presentation of the next set of stimuli. The Presentation software saved a logfile and a text file for each subject. The logfile recorded timestamps of events during the experiment and the text file stored the questions presented and the answers given to them.

2.3 Stimuli, complexity analysis, and behavioral testing

The experimental conditions differed from each other according to the level of difficulty and intelligibility of continuous speech presented. The audio samples were processed using Audacity software. The sound samples that were recorded with sampling rate was 44.1 kHz were cut to the appropriate length and modified to minimize excess noises and to background noise. The beginnings and endings of the samples were harmonized so that there were 2 seconds of silence at the beginning and at the end. The samples were saved as 16-bit WAV files.

In order to make comparisons between different levels of speech intelligibility, vocoded speech stimuli were generated. The audio clips from radio programs made for children were vocoded and vocoded stimuli were edited into mono sounds using Audicity software. They were cut into three parts, the middle of which has an audible target word. For the 1st and 3rd parts noise vocoding was done using 4 channels and speech rotation with a custom Matlab script made by Peelle et al. (2013). Here, the frequency range was divided into 4 logarithmically spaced channels, for each of which the lowpass filter with a cutoff of 30 Hz was applied. After that, the script modulated white noise to this envelope and filtered it again before recombining 4 channels. In the rotation operation, the spectra of

low pass filtered speech were inverted. The 2nd parts of the speech samples were vocoded in Matlab by using 12 channels which leaves the word very understandable but makes it less distinctive from other stimuli. This vocoding makes speech unintelligible as the amount of spectral detail present representing linguistic information is low but overall amplitude envelope is preserved causing speech rhythm properties to be preserved. Noise vocoding has been seen as an effective method to manipulate the spectral and temporal detail while preserving the amplitude envelope (Shannon et al., 1995). Phonetic information is considered to be transmitted in speech edited this way, primarily through temporal envelope rather than spectral cues (Roberts et al., 2011). The technique has been used widely in brain imaging studies and it allows to examine brain responses separately in terms of the rhythmic and linguistic properties of speech stimuli.

To ensure that the easy and the difficult conditions differed from each other linguistically for comparison between speech difficulty levels, they were transcribed and their linguistic complexity was analyzed. In the examination of the complexity of language, the focus was on grammatical and lexical complexity. The measures used to examine grammatical complexity were clause lengths, AS-unit lengths, and the number of dependent clauses per clause. The number of unique words within the first 90 words of each clip, the number of syllables in each word, and the prevalence of each word in Finnish language were used to measure lexical complexity. Statistical analyses were performed in IBM SPSS Statistics version 26 using Mann-Whitney U-test. Statistical significance was set to $p = 0.05$.

The clauses were statistically significantly longer in the difficult clips (mean rank = 169,51) than in the easy clips (mean rank = 107,75), U = 3783,00, z= -6,363, p < 0,001. AS-units (the Analysis of Speech Unit), which according to Foster et al. (2000) are the smallest possible independent units, i.e. independent clauses or sub-clausal units, as well as subordinary clauses associated with either, were statistically significantly longer amongst the difficult clips (mean rank = 103,48) than amongst the easy clips (mean rank = 73,63), U = 2031,00, z = -3,810, p < 0,001. Instead, there was no significant difference in number of dependent clauses per clauses between the easy and the difficult clips.

There were statistically significantly more syllables (mean rank = 681,67) in the words amongst the difficult clips than in the words amongst the easy clips (mean rank = 609,01), U = 179254,50, z = -3,639, p < 0,001. In contrast, the other measures of lexical complexity (prevalence of words and number of unique words within the first 90 words of the transcribe) did not statistically differ significantly between easy and difficult clips.

In addition to the complexity analyses the syllable frequency of each clip were calculated to make sure that differences were not a result of different speaking rates. The frequency

did not differ statistically significantly between difficult (mean rank = 6,80) and easy (mean rank = 4,20) conditions, U = 6,00; z = -1,358; p = 0,175.

The linguistic difficulty of the experiment was tested behaviorally. The subjects (n = 10) were aged between 21 and 28 (average 24,2). One test subject was left-handed, the others were right-handed. At this stage, we measured how well the subjects could answer the questions, asked them to rate the intelligibility of each section on a scale (1–4) and collected open feedback on the test design. On average, subjects answered 4,5/5 questions correctly for vocoded stimuli, 8,5/10 for easy stimuli and 8,6/10 for difficult stimuli.

How understandable the subjects found the easy and the difficult sound samples presented were examined using sum variables (0-20). The median was 19 both in the difficult stimuli (average 19,3) and in the easy stimuli (average 18,5). A statistically significant difference was found between difficult and easy stimuli evaluations (Wilcoxon signed rank test, median of differences in estimates 1; z = 2,333; p = 0,020).

2.4 Processing and analyzing data

2.4.1 MEG data preprocessing and noise reduction

MEG measurements often have contributions from external interference, including noise from the sensors (Clarke et al., 2020). Because the goal was to obtain an accurate representation of brain signals, noise suppression was conducted. MEG data preprocessing and noise reduction was done with MNE-Python (Gramfort et al., 2013). As MEG has limited spatial but good temporal resolution, the use of noise suppression methods that operate in the temporal domain has been the most recommended ones (Clarke et al., 2020). Therefore, two temporal projection noise suppression algorithms for MEG measurements were used in this research.

Firstly, the Oversampled Temporal Projection (OTP) method was used. According to creators of this method OTP requires only that signals are spatially oversampled which, according to them, is often realized in MEG systems (Larson & Taulu, 2018). Larson and Taulu (2018) consider this method, which suppresses uncorrelated sensor noise, to be effective to optimize noise suppression performance while controlling for spatial bias of the signal of interest.

Secondly, another noise reducing method, Temporal Signal Space Separation (tSSS) was used. This commonly used method including the temporal extension of SSS removes artifacts by utilizing time information in addition to Maxwell's equations (Larson & Taulu, 2018). It suppresses

signals from sources that are too close to the sensors for the associated signals to be classified as internal or external assuming that the brain signals and artifact signals are uncorrelated in time (Clarke et al., 2020). tSSS st_correlation parameter was selected to be .98. This parameter is used to determine which MEG signals that are not clearly inside or outside the sensors are artifacts in nature. This is done by testing whether these signals correlate with the signals inside the sensors. Reducing the value of the parameter increases the likelihood that the signals from the interface will be interpreted as an artifact and removed from the remaining data to be analyzed. Clarke et al. (2020) showed that using OTP and tSSS together results in improved signal-to-noise ratio and better detectability of high-frequency oscillations.

MEG signals are contaminated by non-neuronal artifacts like cardiac artefacts (CAs) and eye movements (Treacher et al., 2021). In this research artefacts were removed by using Independent Component Analysis (ICA). One of the MNE-Python software's ICA algorithm, Fastica was selected. ICA method has been shown to be useful when solving so called blind source separation (BSS) problem where unknown signal sources are mixed (Westad & Kermit, 2009). ICA assumes that component signals are statistically independent (Mäkelä et al., 2022). In other words, it assumes that these statistically independent signals origins from different physical processes (Stone, 2002). Cardiac signal detection was done from MEG data, eye movement detection from EOG channel except in the case of two subjects from whom EOG data was not obtained. From them, the detection of the eye movement artifact was made from a frontal magnetometer.

2.4.2 Channel data, quality review, and previewing effects

Correct time intervals were extracted from the data and spectral estimations i.e., estimating the power in different frequency ranges were done by using Welch's method in Matlab. In Welsch's method Power spectral density (PSD) estimation is done by splitting signal into overlapping segments and then applying a window function on segments. After that squared magnitude of Discrete Fourier Transform (DFT) is calculated for segments and individual PSDs are averaged, leading to spectral estimates of the MEG signals. Power spectra estimation methods can be divided into two categories: parametric and non-parametric where the former estimate parameters in a pre-defined model while the latter do not make any assumptions about data (Martini, 2022). Welch's method is a popular non-parametric method (Thomas & John, 2015). Reduced noise due to the averaging is an advantage of the method, whereas its weaknesses include lowered spectral resolution.

2.4.3 Solution to forward-problem

The forward solution was computed. Forward modeling refers to the estimation of MEG field distribution for a known source. To reduce computational costs, a single layer watershed BEM surfaces was computed using the FreeSurfer watershed algorithm. The option ico-4 was used to get approximately 2500 source points per hemisphere. Each subject's data were registered to an average cortical surface space (Freesurfer 5.1.0, "fsaverage" brain)

2.4.4 Calculating a cross-spectral density (CSD) matrix

In the present study, Dynamic Imaging of Coherent Sources (DISC) was used to obtain the source-level estimates of neural activity (Gross et al., 2001). The source-level estimation with DISC is based on a cross-spectral density (DCD) matrix. The CSD is constructed by computing the cross products of the frequency representations between all sensor-to-sensor pairs. Cross spectrum describes the correlation between two signals in the frequency domain.

Welch's method, implemented in Matlab, was used.

2.4.5 Brain level estimation

The source-level estimates of spectral power and coherence between speech and neural signals were calculated at the parcel level, i.e., using anatomical partitioning of the cortex. Spectral coherence is a commonly used method to quantify phase locking between signals and it can said to be computationally fast ja robust against noise (Lowet et al., 2016). The calculated power and coherence covered the range from 2 to 45 Hz.

In this study gyral based Freesurfer atlas called the Desikan-Killiany atlas was used as the cortical parcellation meaning that for each participant the cortex was parcellated into 68 subregions. Regions of interest (ROI) examined within the present study are presented in Fig.1.
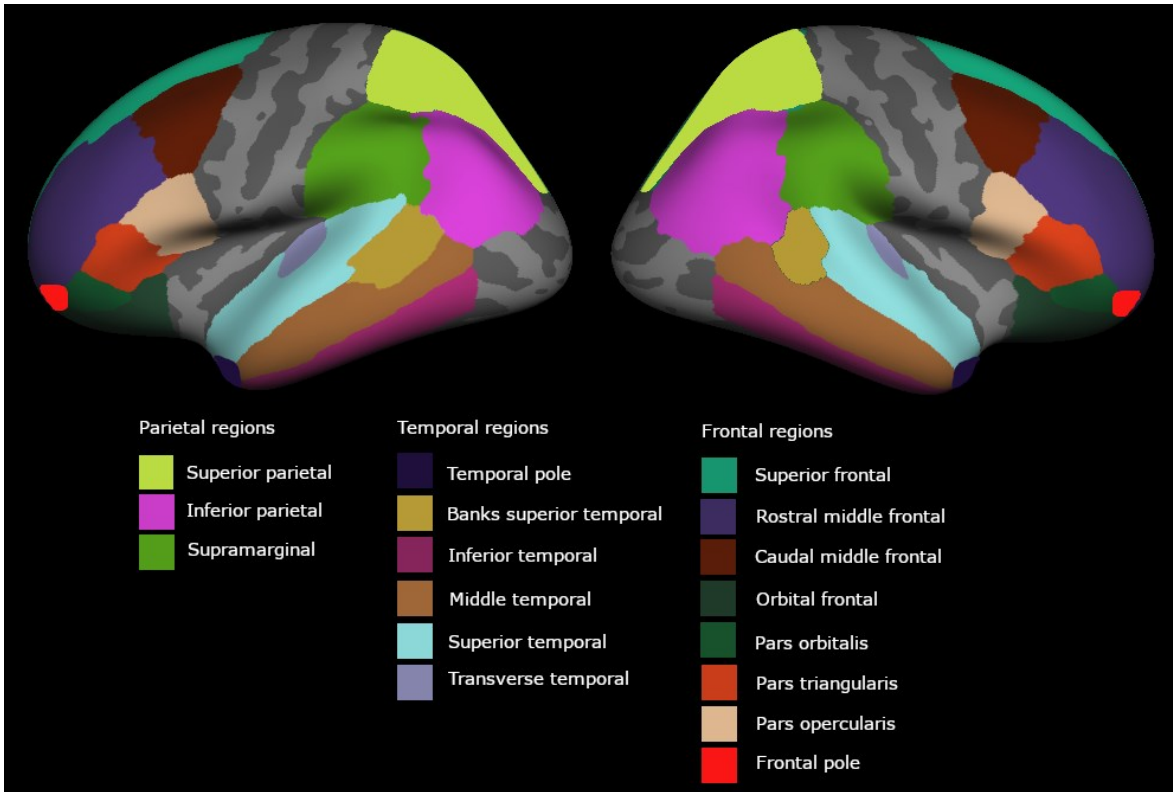
Fig 1. Regions of Interest (ROI) based on the Desikan-Killiany atlas.

2.4.6 Parameterizing neural power spectra

Electrophysiological brain data is typically not mainly rhythmic but contains a mixture of oscillatory and aperiodic activity (Donoghue & Watrous, 2022). Therefore, to be able to examine aperiodic brain activity and oscillations separately, the Fitting oscillations and one over f algorithm (FOOOF) was used to all parcels power spectra. FOOOF is frequently used method to separate aperiodic and periodic components. It models periodic components and an aperiodic component and outputs the corresponding parameters while the second most used method IRASA only separates them (Gerster et al., 2022). Donoghue et al. introduced the FOOOF algorithm as a method to parameterize neural power spectra as a combination of an aperiodic component and putative periodic oscillatory peaks (Donoghue et al., 2020). According to them their method is more robust and generalizable than previous ones (Donoghue & Watrous, 2022). The FOOOF algorithm does not require specification of narrowband oscillation frequencies as it identifies oscillations based on the power above the aperiodic component (Donoghue et al., 2020). Another strength of this algorithm is that it gives a measure of aperiodic components which allows comparison of aperiodic components both within and between subjects (Donoghue et al., 2020).

Using this algorithm in this study, both 1/f offset and exponent estimates were obtained.

2.4.7 Statistical analysis

The aperiodic signal, calculated based on offset and exponent parameters, was removed from the original power spectrum. The differences between canonical oscillations (theta, 4-7 Hz; alpha, 8-13 Hz; beta, 13-25 Hz; gamma, 25-45 Hz) were calculated between different conditions (easy vs. difficult, and easy vs. vocoded speech). The hertz boundaries used with oscillations vary somewhat in the research literature. It should be noted that in this study, 25-35 Hz oscillations are called gamma oscillations, while in some studies they are called high-beta oscillations.

We also examined the differences between the conditions for the offset and exponent parameters of aperiodic activity and audio-MEG coherence averaged at 2-12 Hz. All statistical tests were conducted using Student's two-tailed t-tests for paired samples and were done in Matlab.

Statistically significant ($p<0.05$) results were stored in files to enable visualization with Freesurfer.

2.4.8 Visualization of results

The Freesurfer brain imaging software package with TKSurfer tool was used for visualization and anatomical division of results.

3. RESULTS

3.1 Periodic brain activity

The first research question examined the differences in rhythmic, periodic brain activation between different speech conditions. The examination was done in the frontal, parietal and temporal regions of the brain and the differences were considered separately in the left and right hemispheres. T-tests were done with power values (tesla²/m²) (Appendix B). Statistically significant differences (statistical significance threshold $p < .05$) were observed in theta-, beta- and gamma-band oscillations. In comparisons of speech difficulty levels, differences were found in 25-35 Hz and 35-45 Hz oscillations, and in comparisons of different speech intelligibility levels in 4-7, 13-20, 15-25, and 25-35 Hz oscillations.
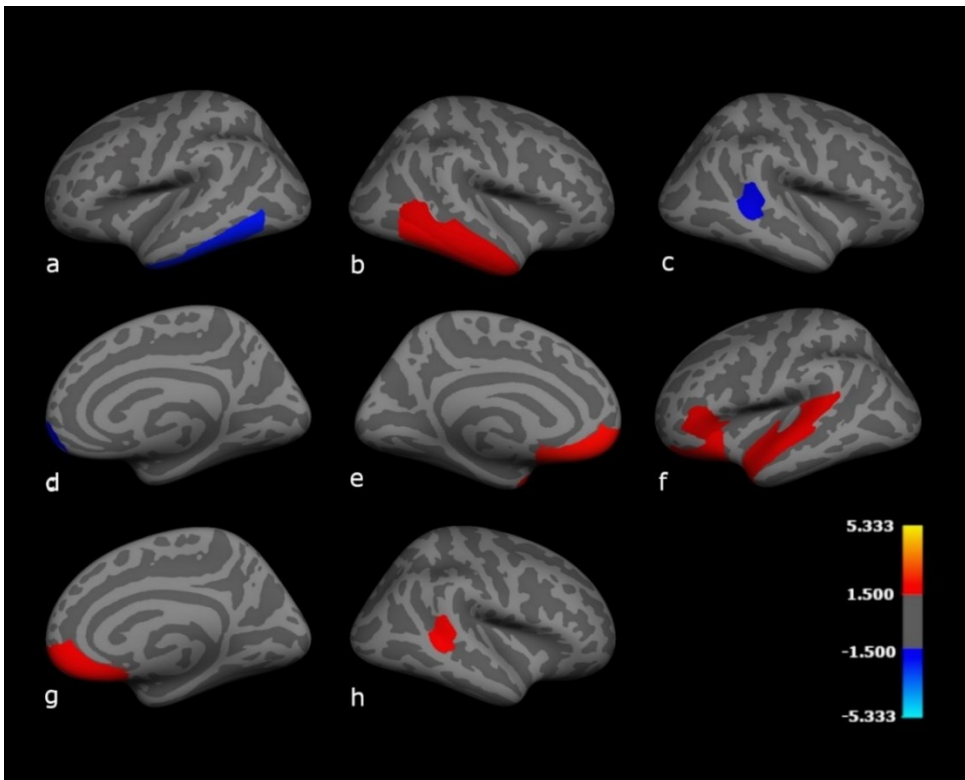
Figure 2. T-tests results for power values. Significant differences in speech perception-related activity depending on experimental condition (p < 0,05; uncorrected) overlaid onto the cortical anatomy. a) easy < vocoded speech, theta oscillations (4-7 Hz), b) easy > vocoded speech, beta oscillations (13-20 and 15-25 Hz), c) difficult < easy speech, beta oscillations (25-35 Hz), d) difficult < easy speech, gamma oscillations (35-45 Hz), e-h) easy > vocoded speech, beta oscillations (25-35 Hz)

When comparing different difficulty levels (easy vs. difficult speech) with each other, significantly more neural activity was observed in the beta-band (25 - 35 Hz) and gamma-band (35 - 45 Hz) in easy than in difficult condition. In beta-band difference was found in the right banks sts and in gamma-oscillations in right frontal pole area.

Also, when making comparisons between different levels of speech intelligibility (easy vs. vocoded), some statistically significant different were found. In the theta-band (4 -7 Hz) statistically significant differences were found when comparing easy and vocoded speech. More neural activity was found for vocoded than easy speech perception in the left inferior temporal region (Fig. 2a). Statistically significant differences were also found in the 13-25 Hz oscillations as there was more neural activity in the easy than in the vocoded setting in the right inferior and middle temporal regions (fig 3b).

Also, more neural activity was detected in the easy than in the vocoded condition in the left lateral and medial orbitofrontal area, pars triangularis, superior temporal area, frontal and

temporal poles (25-35 Hz) (fig. 3e & f) and right banks sts., medial orbitofrontal area and frontal pole (25-35 Hz) (fig. 3g & h).

No statistically significant differences were found neither in comparisons of speech difficulty nor speech intelligibility levels in the delta (1-4 Hz) or alfa (8-13 Hz) oscillations. As mentioned above oscillations over 45 Hz were not examined.

3.2 Aperiodic brain activity

The second research question examined at the differences in arrhythmic, aperiodic brain activity between different levels of speech difficulty (easy vs. difficult) and different levels of speech intelligibility (easy vs. vocoded). Also, this examination was done in the frontal, parietal and temporal regions of the brain. The differences were considered separately in the left and right hemispheres. T-tests were done with the exponent and the offset values (Appendix C). Statistically significant differences (statistical significance threshold $p < .05$) were observed only when comparing different levels of speech intelligibility. These findings were obtained both in the exponent and the offset values.

Firstly, exponent values were significantly larger for the easy than in the vocoded speech condition in the left rostral middlefrontal and frontalpole regions (Fig3a-b). Secondly, offset values were statistically significantly larger for the easy than in the vocoded speech setting in the left pars orbitalis, rostral middlefrontal and frontalpole regions (Fig.3c-d) and right rostral middlefrontal region (Fig.3e).

Comparisons of different speech difficulty levels did not reveal statistically significant differences in either measure of aperiodic brain activity.
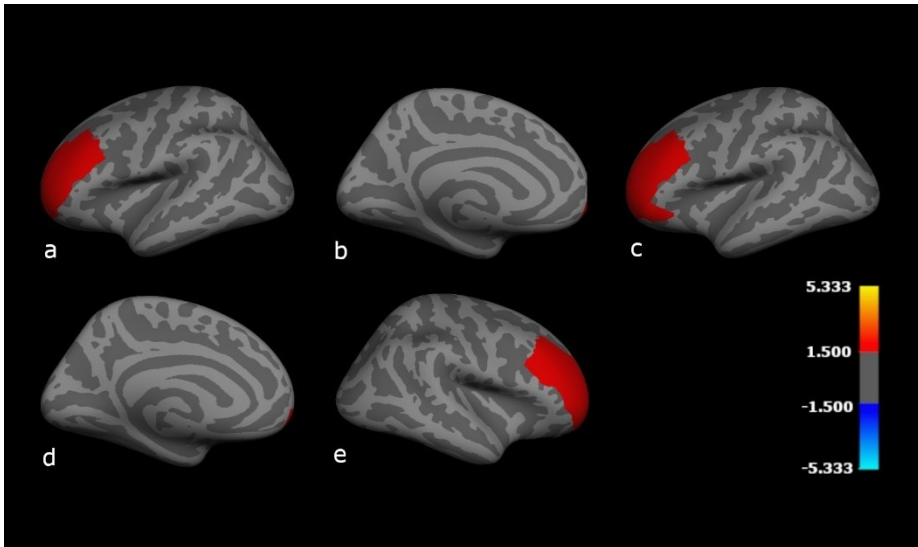
Figure 3 T-tests results for the exponent and offset values. Significant differences in speech perception-related activity depending on speech condition (p < 0.05, uncorrected) overlaid onto the cortical anatomy. a-b) exponent, easy > vocoded speech, left, c-d) offset, easy > vocoded speech, left, e) offset, easy > vocoded speech, right

3.3 Phase-locking to the envelope of speech signal

In the third neural activation phenomenon examined, in phase-locking to the envelope of speech signal, some statistically significant differences were found in comparisons of both speech difficulty levels and speech intelligibility levels. Audio–MEG coherence values (Appendix D) were compared across conditions using Student's two-tailed t-test (statistical significance threshold $p < .05$). Statistically significant differences were observed both in the left and right hemispheres.

When comparing different levels of speech difficulty with each other, audio-MEG coherence was stronger for difficult than in easy condition in the left medial orbitofrontal, superior parietal and frontal pole regions (Fig.4a-b) and right caudal middlefrontal and pars opercularis regions (Fig.4c).

In turn, when making comparisons between different levels of speech intelligibility, coherence was stronger in the easy than in the vocoded setting in right supramarginal and temporalpole regions (Fig.4d-e).
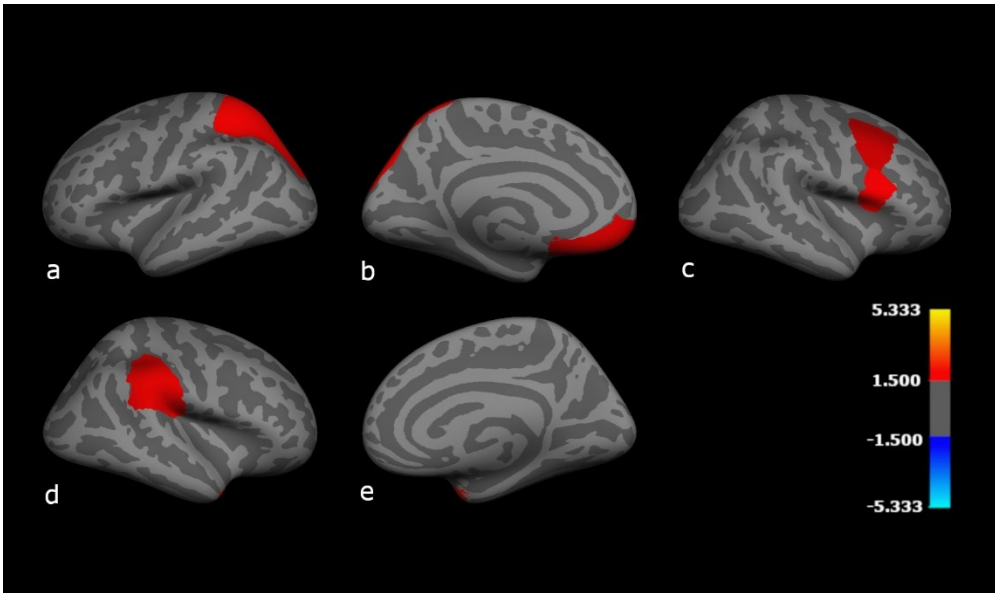
Figure 4 T-tests results for audio-MEG coherence values. Significant differences in speech perception-related activity depending on the condition (p < 0.05, uncorrected) overlaid onto the cortical anatomy. a-b) audio-MEG coherence, difficult > easy, left medial orbitofrontal, superiorparietal and frontalpole regions, c) audio-MEG coherence, difficult > easy, right caudal middlefrontal and pars opercularis region, d-e) audio-MEG coherence, easy > vocoded, right supramarginal and temporalpole regions

4. DISCUSSION

The purpose of this study was to examine whether differences can be found in rhythmic and arrhythmic brain activity and in the locking of brain activity to speech stimuli when subjects listen to speech stimuli of varying degrees of difficulty and intelligibility. The aim was to gain a better understanding of the significance of these different phenomena in terms of the level of difficulty and intelligibility of the speech to be perceived. For this purpose, the three above mentioned brain activity measures were quantified with a test design in which the subjects were presented with continuous speech of varying difficulty and intelligibility.

The statistical comparisons included the frontal, temporal and parietal brain regions known to be central in speech processing. Statistical testing was used to identify possible differences in spectral power, coherence, exponent, and offset values. The study aimed to correct the shortcomings of the previous studies by using natural speech as stimuli, in addition to which the different levels of speech difficulty and intelligibility were taken into account. In this study, speech intelligibility refers to the extent to which spectral and temporal detail of stimuli has been manipulated while preserving the amplitude envelope whereas speech difficulty level refers to linguistic complexity of stimuli when the focus is on grammatical and lexical complexity.

The research questions were based on observations made in previous studies, according to which many mental functions have been linked to both periodic and aperiodic brain activity. For different oscillations, especially delta and theta oscillations, such associations have been found for speech processing. In addition, although there is little previous research data on arrhythmic brain activity, it has been suggested that 1/f activity plays an important role in higher order cognition like language processing and that 1/f activity is a predictor of N400 effects of successful lexical prediction. Thirdly, also cortical activity locking on to the envelope of speech signal was included in the phenomena under consideration.

Although no actual hypotheses were presented related to the research questions, cautious assumptions were made that easy speech might differ from difficult (speech difficulty) and easy speech from vocoded (speech intelligibility), generating modulations in all examined neural mechanisms in different areas of the brain. These expectations were not fully met, although some statistically significant differences were found.

4.1 Speech difficulty

When comparing speech content of different difficulty levels (easy vs. difficult), differences were found between periodic brain function and in phase-locking (audio-med coherence). On the other hand, there were no differences between the different levels of speech difficulty in the offset and exponent values measuring aperiodic brain activity.

Higher spectral power was associated with easy stimuli compared to difficult speech both in beta and gamma oscillations. Previous studies have made some observations on the association between power spectrum in beta-band and processes that are considered important for speech perception, i.e. converting continuous variations in acoustics into discrete units (Bidelman, 2015) and discriminating tones (Ou & Law, 2019). Since in the latter study higher beta oscillation power was associated with higher tone separation, the result obtained in this study of higher beta-band power values when receiving easier than difficult speech could be interpreted cautiously so that when listening to easier speech, speech tone separation would be more successful. However, more research is needed on this possible link.

More neural activity was observed in the beta-band in the easy than in difficult condition in the right banks of the superior temporal sulcus (i.e., the posterior aspect of the superior temporal sulcus, rpSTS). Although the importance of the left hemisphere in speech processing has traditionally been emphasized, it is now known that also the right hemisphere participates in linguistic processes. Therefore, the above result cannot be completely surprising although some differences in beta-bands might have been expected to be found in the left hemisphere as well. Also, the fact that the above

result was obtained specifically in the temporal regions of the brain cannot be considered surprising, since temporal areas are considered central to the perception of speech, which is partly due to the fact that the primary auditory areas are located bilaterally in them.

In addition to beta oscillations, higher spectral power in easy than in difficult condition was found in the fastest brain waves, 35-45 Hz referring as gamma oscillations in this study. This is not a completely surprising finding in that fastest oscillations are also believed to play roles in cognitive processes and as bursts of gamma oscillations have been observed in response to auditory stimulation. Previous studies have found links between the fastest brain waves and the characteristics of the stimuli presented. In this study, differences in the fastest oscillations were observed in the right frontal pole region. This prefrontal region is considered to be related to cognitive function and control of behavior. However, the result obtained here, according to which the power of fast oscillations was greater in easier speech than in more difficult speech, could be considered somewhat surprising and even contrary to expectations. It has previously been observed, for example, that moving visual stimuli generates higher frequency gamma oscillations than stationary stimuli (Muthukumaraswamy & Singh, 2013; Swettenham et al., 2009). Whether the result obtained here is due to a small sample size or some still unknown speech processing phenomenon related to these fastest oscillations, which is reflected in a decrease in the power of gamma oscillations as speech difficulty levels become more difficult, cannot be said with certainty here and the phenomenon requires further study.

The oscillations examined above have been considered to indicate the rhythmicity of brain activity, in addition to which speech has been presented to be characterized by rhythmicity independently of language and the regularity of the rhythm of speech is suggested to play a role in speech comprehension. Attempts have been made to investigate these findings by looking for connections between these two rhythmicities, with the aim of determining whether brain activity follows the rhythm of the received speech stimuli. In this context, brain activity's entrainment to temporal structures of the speech have been discovered. In this study, some differences in phase-locking were found when comparing speech of different difficulty levels. In this comparison, a more difficult speech stimulus resulted in higher audio-meg coherence than an easy one in left medial orbitofrontal, superior parietal, and frontal pole and in right caudal middlefrontal and pars opercularis areas. The result can be considered somewhat contrary to expectations since phase-locking is assumed to be an important phenomenon for speech understanding. However, there is little previous research on the subject, and on the other hand, there has also been criticism of the locking between speech and brain function. There is no consensus on the temporal characteristics of speech, it has been suggested that the rhythm of speech depends on the speech stimuli presented, and on the other hand, it has also been suggested that rhythmic structure is not typical for natural connected speech, which was

examined in this study. It may also be that the level of difficulty of speech referring to linguistic, especially grammatical and lexical, complexity of stimuli, examined in this comparison is not related to the phenomenon of locking, which could be due, for example, to the fact that such variation in speech difficulty does not alter its rhythmic characteristics. In this case, the result obtained here may be random and due to, for example, a small sample size.

However, when comparing difficult and easy speech in terms of phase-locking, such statistically significant results were obtained in areas of the brain that can be considered interesting for speech processing. The prefrontal cortex is central in the executive control of information processing and behavioral expression, and its part, the orbitofrontal cortex, has been shown to participate in the executive control and behavioral expression of information processing (Hooker & Knight, 2006). The superior parietal region, on the other hand, is known to be significantly involved in different neurological functions, especially in visual and attention but it has also shown to have connections with the language network in left inferior frontal gyrus (Alahmadi, 2021). In turn, caudal middlefrontal area is interesting, because middle frontal gyrus has been found to make a distinction between intelligible and non-intelligible speech in such a way that it is able to differentiate natural and synthetic speech (Benson et al., 2001). Furthermore, as mentioned earlier, the frontal pole is considered to be related to cognitive function and control of behavior, while the pars opercularis is part of the Broca area, known as central to speech.

4.2 Speech intelligibility

While when comparing different levels of speech difficulty, statistically significant differences were found only in periodic brain activity and the locking of brain activity into speech stimuli, comparing different levels of speech intelligibility (easy vs. vocoded) also produced statistically significant differences in aperiodic brain function.

With regard to periodic brain function, while comparing speech difficulty levels produced greater powers in easy speech than in difficult speech, both results were obtained in terms of speech intelligibility in this review (fig. 2). In theta oscillations, the difference differed in direction from other findings, as in vocoded speech the power was greater than in easy speech in the left inferior temporal area. Other statistically significant results were reversed and related to beta oscillations. In them, power was greater in more intellectual speech in beta oscillations (13-35 Hz) in right inferior temporal and middle temporal areas and right banks sts, medial orbitofrontal and frontal pole areas and in left lateral orbitofrontal, medial orbitofrontal, past triangularis, superior temporal, frontal pole and temporal pole regions. Despite the partly contradictory results between theta and beta, all statistically significant results were found in brain regions of interest for speech processing.

The fact that statistically significant results were obtained from theta and beta oscillations is not contrary to expectations, as links have been found with both in terms of speech processing. A burst of theta activity is known to occur when processing auditory information (Picton & Picton, 2010) in addition to which cognitive load has been found to increase power of theta oscillations (Snipes et al., 2022). In this study, theta oscillation powers were higher in vocoded, i.e. more difficult to understand speech than in easy speech, which is consistent with the above observation. In this case, greater power in beta oscillations may be the result of the greater cognitive effort required for vocoded, i.e. less intelligibility, speech. The finding underscores the possible importance of theta oscillations for speech understanding, but the phenomenon requires more research with a larger sample size.

In contrast, in beta oscillations, lower power in less intelligibility setting was somewhat contrary to expectations, as previous studies have found correlation between activity in beta oscillations and the ability to convert continuous variations in acoustics into discrete units, which is important in the perception of speech (Bidelman, 2015). On the other hand, beta oscillations have been shown to reflect individual differences in phonological discrimination sensitivity (Ou & Law, 2019). Thus, in this study, where the sample size was relatively small, the phenomenon found may be explained by the sensitivity characteristics of the studied individuals in terms of phonological processing, and the result may not so much reflect the role of beta oscillations in speech understanding. Studying the phenomenon with a larger sample size could provide additional information about the true nature of the phenomenon.

The absence of delta oscillations in the significant results of this study was somewhat contrary to expectations, as delta oscillations have also been found in previous studies to play a role in speech processing and the intelligibility of speech stimuli. On the other hand, delta oscillations have been suggested to play a role different from theta oscillations in speech perception, with delta oscillations being suggested to be more related to perceived acoustic rhythm rather than the phonemic information whereas theta oscillations would be more related to syllabic-level acoustic features (Ding & Simon, 2014). Thus, it may be that in this study, the lack of statistically significant differences in the comparison between in delta power vocoded and easy speech is due to the fact that delta power is relevant for acoustic rhythm, and in vocoding, speech rhythm properties are preserved. In this case, different roles for delta and theta oscillations in speech comprehension are assumed.

Periodic brain activity has already been studied quite extensively, while aperiodic activity has often been ignored or considered as noise. However, there is evidence that aperiodic brain activity also has connections with various phenomena, for example excitation-inhibition balance and various diseases and the study of the aperiodic component is justified. In this study, a comparison of

different speech intelligibility (easy vs. vocoded speech) found statistically significant results also in aperiodic brain function. Results were obtained for both variables measuring aperiodic brain activity, exponent and offset, and in these the results were similar. All the findings indicated that aperiodic brain activity was greater in easier than in vocoded speech. When examining exponent variables reflecting the pattern of aperiodic power across frequencies, higher values were obtained in more understandable speech in left rostral middlefrontal and frontal pole regions, whereas in offset variable reflecting uniform shift of power across frequencies values were higher in left pars orbitalis, rostral middlefrontal and frontalpole regions as well as right rostral middlefrontal regions. All these areas of the brain have already been identified above as interesting for speech perception. The results were similar to those obtained for beta oscillations in periodic brain activity, where powers were greater in easy speech than in vocoded speech. Based on the results of this study, one could make very cautious assumptions that both periodic and aperiodic brain activity are in some sense relevant for speech intelligibility. What exactly their role is in speech understanding and what role different oscillations such as beta and theta play in periodic brain activity is not clear from the results of this study.

Also, in the third neural activation phenomenon examined, in phase-locking to the envelope of speech signal, some statistically significant differences were found in comparisons of speech intelligibility levels (easy vs. vocoded speech). Contrary to expectations, however, the result was opposite of when comparing speech difficultness levels to each other as phase-locking was stronger in difficult than in easy speech, whereas in intelligibility locking was stronger in easy than in vocoded speech. When comparing audio-meg coherence between easy and vocoded speech, stronger phase-locking was found in right supramarginal and temporal pole regions. For both brain regions, previous observations have been made about their role in some phenomena relevant to speech processing. In general, the supramarginal gyrus has been associated with non-linguistic processes, but interestingly for the findings of this study, it has also been bilaterally linked with phonological decision-making. Specifically, it has been shown to be activated when a person makes decisions about the meaning of words (Hartwigsen et al., 2010). However, in this study, the finding was made only for the right side and the result was not obtained in comparisons other than between the easy and the vocoded setting. The result obtained in this study for the temporal pole was also not surprising, since previous studies have indicated that this area plays a role in the perception of speech. Mazoyer et al. (1993) found that whereas listening to an unknown language activated both temporal poles symmetrically, listening to meaningful stories more than doubled this activation level and resulted in a proportionally higher left-side activation. Also, the result obtained in this study, the greater gamma-band power in the left temporal pole for intelligible speech may suggest that speech intelligibility is associated with higher activation specifically in the left temporal pole area.

4.3 Summary of the results

The results of this study reinforced the notion that speech perception is diverse in its neural basis. Statistically significant differences were found in all three types of brain activity measures when comparing the perception of speech stimuli with different levels of difficulty and intelligibility. However, the results were partly contrary to expectations and partly contradictory.

For example, the absence of considered delta oscillations, which are important for speech understanding, in statistically significant results did not meet expectations. There were also no significant differences in alpha oscillations, either in speech difficultness or in intelligibility comparisons, although for example Ala et al. (2022) have suggested that alpha power in continuous speech can correlate with how well speech is perceived. However, this study did not provide evidence for this hypothesis.

Conflicting were the results in which intelligibility and difficultness appeared to be opposite in phase-locking. Phase-locking was greater in difficult than in easy speech and in easy than in vocoded speech, although one might have expected that more difficult and vocoded speech would have produced similar results when comparing to easy speech.

In contrast, the results in periodic brain activity were mostly similar, with the exception of theta oscillations, which in intelligibility differed from other comparisons. In general, oscillation powers were larger in easy speech than in difficult or vocoded speech, i.e. as the speech stimulus to be received became more cognitive effort-intensive, oscillation powers decreased – with the exception of theta oscillations, where the power was higher in vocoded speech requiring more effort.

The results of this study were not strong enough to make definite assumptions about the significance of periodic and periodic brain activity or the phenomenon of brain locking in speech reception, or about the more precise nature of these phenomena. The results can therefore be considered more as raising questions for further research and underlining the diversity of brain activity involved in processing speech reception.

4.4 Advantages and limitations of the study and topics for further research

Today it is known that the nervous basis of speech is a complex phenomenon. Localization theories limited to individual brain regions have been abandoned and the focus has been on the neural networks involved in the production and perception of speech. This challenges the author of the research. This study took into account this involvement of several brain regions in speech perception by conducting examination in the broad frontal, parietal and temporal cortex regions, as they are

known, in the light of previous research, to participate in linguistic functions and the processing of speech perception. The strength of the study can also be considered the fact that MEG as the measurement method used in it provided high temporal resolution.

In addition to taking into account large areas of the brain, the stimuli used are important when studying speech perception. In previous research, the neural basis of speech has been largely examined using individual speech stimuli. Thus, the strength of this study was that instead of isolated sentences and read-aloud texts it used natural continuous speech, which better corresponds to the stimuli perceived in real speech situations. Behavioral testing was carried out to ensure the functionality of the test design. In addition, a statistical analysis of the difficulty level of the speech samples was done, which made it possible to ensure that the different speech settings differed statistically significantly from each other in terms of difficulty.

However, the speech stimuli used were selected from the already existing audio material. In terms of different linguistic characteristics, a more controlled stimuli would be obtained if speech stimuli made for each study were used. Furthermore, the sample size of the study was quite small, which may have limited results obtained. With a larger sample size, more statistically significant results might have emerged. In addition, some of the results obtained may be caused by statistical error and may not be explainable by existing phenomena.

Although this study did not provide clear and consistent results that could be used to explain the neural processing of speech of different levels of difficulty and intelligibility, results underscore the need for additional work that considers larger sample size and more controlled stimuli. Using a larger sample size could reveal more statistically significant results. It would be a good idea to prepare the speech stimuli to be used for research instead of using ready-made speech material, the linguistic features of which cannot thus be fully controlled. A larger sample size could provide more information on, for example, the potentially different roles of delta- and theta-band oscillations in speech understanding, while in this study statistically significant results for delta-band oscillations were missed.

In this study, the subjects listened to their own mother language, Finnish. In future studies, it would be interesting to use other languages as stimuli in order to highlight the potential special features of speech processing caused by different languages in brain activation data. Also, the multisensory nature of speech perception would be an interesting topic for further research from the perspective of all neural phenomena studied here. Natural speech is often multimodal and contains visual information in the form of visual speech including lip-reading. In this study, speech perception was studied only in terms of auditory processing and the stimuli used included only speech sound. Incorporating the speaker's facial expressions and mouth movements into future studies would bring

the study even closer to the natural speech situations and provide information on the processing of multimodal speech stimuli.

REFERENCES

Ai, L., & Ro, T. (2016). Alpha oscillations and desynchronizations facilitate visual-tactile multisensory integration. *Journal of Vision*, *16*(12), 1189. https://doi.org/10.1167/16.12.1189

Ala, T. S., Alickovic, E., Cabrera, A. F., Whitmer, W. M., Hadley, L. V., Rank, M. L., Lunner, T., & Graversen, C. (2022). Alpha Oscillations During Effortful Continuous Speech: From Scalp EEG to Ear-EEG. *IEEE Transactions on Bio-Medical Engineering*, *PP*. https://doi.org/10.1109/TBME.2022.3214428

Alahmadi, A. A. S. (2021). Investigating the sub-regions of the superior parietal cortex using functional magnetic resonance imaging connectivity. *Insights into Imaging*, *12*(1), 47. https://doi.org/10.1186/s13244-021-00993-9

Alexandrou, A. M., Saarinen, T., Kujala, J., & Salmelin, R. (2016). A multimodal spectral approach to characterize rhythm in natural speech. *The Journal of the Acoustical Society of America*, *139*(1), 215–226. https://doi.org/10.1121/1.4939496

Alexandrou, A. M., Saarinen, T., Kujala, J., & Salmelin, R. (2020). Cortical entrainment: What we can learn from studying naturalistic speech perception. *Language, Cognition and Neuroscience*, *35*(6), 681–693. https://doi.org/10.1080/23273798.2018.1518534

Arana, L., Melcón, M., Kessel, D., Hoyos, S., Albert, J., Carretié, L., & Capilla, A. (2022). Suppression of alpha-band power underlies exogenous attention to emotional distractors. *Psychophysiology*, *59*(9), e14051. https://doi.org/10.1111/psyp.14051

Assaneo, M. F., Rimmele, J. M., Orpella, J., Ripollés, P., de Diego-Balaguer, R., & Poeppel, D. (2019). The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms. *Frontiers in Integrative Neuroscience*, *13*. https://www.frontiersin.org/articles/10.3389/fnint.2019.00028

Bäckström, T., Räsänen, O., Zewoudie, A., Zarazaga, P. P., Koivusalo, L., Das, S., Mellado, E. G., Mariem Bouafif Mansali, & Ramos, D. (2022). *Introduction to Speech Processing: 2nd Edition*. https://doi.org/10.5281/ZENODO.6821775

Barone, J., & Rossiter, H. E. (2021). Understanding the role of sensorimotor beta oscillations. *Frontiers in Systems Neuroscience, 15,* Article 655886. https://doi.org/10.3389/fnsys.2021.655886

Başar, E., Tülay, E., & Güntekin, B. (2015). Multiple gamma oscillations in the brain: A new strategy to differentiate functional correlates and P300 dynamics. *International Journal of Psychophysiology*, *95*(3), 406–420. https://doi.org/10.1016/j.ijpsycho.2015.01.013

Bastiaansen, M., Berberyan, H., Stekelenburg, J. J., Schoffelen, J. M., & Vroomen, J. (2020). Are alpha oscillations instrumental in multisensory synchrony perception? *Brain Research*, *1734*, 146744. https://doi.org/10.1016/j.brainres.2020.146744

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language*, *78*(3), 364–396. https://doi.org/10.1006/brln.2001.2484

Bidelman, G. M. (2015). Induced neural beta oscillations predict categorical speech perception abilities. *Brain and Language*, *141*, 62–69. https://doi.org/10.1016/j.bandl.2014.11.003

Billig, A. J., Herrmann, B., Rhone, A. E., Gander, P. E., Nourski, K. V., Snoad, B. F., Kovach, C. K., Kawasaki, H., Howard, M. A., & Johnsrude, I. S. (2019). A Sound-Sensitive Source of Alpha Oscillations in Human Non-Primary Auditory Cortex. *Journal of Neuroscience*, *39*(44), 8679–8689. https://doi.org/10.1523/JNEUROSCI.0696-19.2019

Binder, J. R. (2015). The Wernicke area. *Neurology*, *85*(24), 2170–2175. https://doi.org/10.1212/WNL.0000000000002219

Bourguignon, M., De Tiège, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., & Jousmäki, V. (2012). The pace of prosodic phrasing couples the listener's

cortex to the reader's voice. *Human Brain Mapping*, *34*(2), 314–326.

https://doi.org/10.1002/hbm.21442

Boustani, S. E., Marre, O., Béhuret, S., Baudot, P., Yger, P., Bal, T., Destexhe, A., & Frégnac, Y.

(2009). Network-State Modulation of Power-Law Frequency-Scaling in Visual Cortical

Neurons. *PLOS Computational Biology*, *5*(9), e1000519.

https://doi.org/10.1371/journal.pcbi.1000519

Brent, M. R. (1999). Speech segmentation and word discovery: A computational perspective.

Trends in Cognitive Sciences, *3*(8), 294–301. https://doi.org/10.1016/s1364-6613(99)01350-

9

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts

Visual Perception. *Journal of Neuroscience*, *29*(24), 7869–7876.

https://doi.org/10.1523/JNEUROSCI.0113-09.2009

Callan, D., Callan, A., Gamez, M., Sato, M., & Kawato, M. (2010). Premotor cortex mediates

perceptual performance. *NeuroImage*, *51*(2), 844–858.

https://doi.org/10.1016/j.neuroimage.2010.02.027

Cantero, J. L., Atienza, M., & Salas, R. M. (2002). Human alpha oscillations in wakefulness,

drowsiness period, and REM sleep: Different electroencephalographic phenomena within

the alpha band. *Neurophysiologie Clinique/Clinical Neurophysiology*, *32*(1), 54–71.

https://doi.org/10.1016/S0987-7053(01)00289-1

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive

crossmodal illusory perception. *Current Biology: CB*, *25*(2), 231–235.

https://doi.org/10.1016/j.cub.2014.11.034

Chevillet, M. A., Jiang, X., Rauschecker, J. P., & Riesenhuber, M. (2013). Automatic Phoneme

Category Selectivity in the Dorsal Auditory Stream. *Journal of Neuroscience*, *33*(12), 5208–

5215. https://doi.org/10.1523/JNEUROSCI.1870-12.2013

Clarke, M., Larson, E., Tavabi, K., & Taulu, S. (2020). Effectively combining temporal projection noise suppression methods in magnetoencephalography. *Journal of Neuroscience Methods*, *341*, 108700. https://doi.org/10.1016/j.jneumeth.2020.108700

Clayton, M. S., Yeung, N., & Cohen Kadosh, R. (2018). The many characters of visual alpha oscillations. *European Journal of Neuroscience*, *48*(7), 2498–2508. https://doi.org/10.1111/ejn.13747

Cross, Z. R., Corcoran, A. W., Schlesewsky, M., Kohler, M. J., & Bornkessel-Schlesewsky, I. (2021). *Oscillatory and aperiodic neural activity jointly predict language learning* (p. 2020.03.10.984971). bioRxiv. https://doi.org/10.1101/2020.03.10.984971

D'Ausilio, A., Craighero, L., & Fadiga, L. (2012). The contribution of the frontal lobe to the perception of speech. *Journal of Neurolinguistics*, *25*(5), 328–335. https://doi.org/10.1016/j.jneuroling.2010.02.003

Dave, S., Brothers, T. A., & Swaab, T. Y. (2018). 1/f neural noise and electrophysiological indices of contextual prediction in aging. *Brain Research*, *1691*, 34–43. https://doi.org/10.1016/j.brainres.2018.04.007

Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiège, X., & Bourguignon, M. (2019). Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. *NeuroImage*, *184*, 201–213. https://doi.org/10.1016/j.neuroimage.2018.09.006

Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Current Biology*, *25*(19), 2457–2465. https://doi.org/10.1016/j.cub.2015.08.030

Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust Cortical Entrainment to the Speech Envelope Relies on the Spectro-temporal Fine Structure. *NeuroImage*, *88*, 41–46. https://doi.org/10.1016/j.neuroimage.2013.10.054

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, *81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, *109*(29), 11854–11859. https://doi.org/10.1073/pnas.1205381109

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*, 311. https://doi.org/10.3389/fnhum.2014.00311

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

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Donoghue, T., & Watrous, A. (2022). *How can we differentiate narrow-band oscillations from aperiodic activity?* PsyArXiv. https://doi.org/10.31234/osf.io/k6nhd

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *111*, 7126–7131. https://doi.org/10.1073/pnas.1318738111

Etard, O., & Reichenbach, T. (2019). Neural Speech Tracking in the Theta and in the Delta Frequency Band Differentially Encode Clarity and Comprehension of Speech in Noise. *The Journal of Neuroscience*, *39*, 1828–18. https://doi.org/10.1523/JNEUROSCI.1828-18.2019

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Franken, M. K., Liu, B. C., & Ostry, D. J. (2022). Towards a somatosensory theory of speech perception. *Journal of Neurophysiology*, *128*(6), 1683–1695. https://doi.org/10.1152/jn.00381.2022

Fred, A. L., Kumar, S. N., Haridhas, A. K., Ghosh, S., Bhuvana, H. P., Sim, W. K. J., Vimalan, V., Givo, F. A. S., Jousmäki, V., Padmanabhan, P., & Gulyás, B. (2022). *A Brief Introduction to Magnetoencephalography (MEG) and Its Clinical Applications*. https://doi.org/10.3390/brainsci12060788

Gao, R., Peterson, E. J., & Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *NeuroImage*, *158*, 70–78. https://doi.org/10.1016/j.neuroimage.2017.06.078

Gerster, M., Waterstraat, G., Litvak, V., Lehnertz, K., Schnitzler, A., Florin, E., Curio, G., & Nikulin, V. (2022). Separating Neural Oscillations from Aperiodic 1/f Activity: Challenges and Recommendations. *Neuroinformatics*, *20*(4), 991–1012. https://doi.org/10.1007/s12021-022-09581-8

Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, *108*(1), 33–56. https://doi.org/10.1037/0033-295X.108.1.33

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*, 511–517. https://doi.org/10.1038/nn.3063

Glanz (Iljina), O., Derix, J., Kaur, R., Schulze-Bonhage, A., Auer, P., Aertsen, A., & Ball, T. (2018). Real-life speech production and perception have a shared premotor-cortical substrate. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-26801-x

Gosseries, O., Demertzi, A., Noirhomme, Q., Tshibanda, J., Boly, M., Op de Beeck, M., Hustinx, R., Maquet, P., Salmon, E., Moonen, G., Luxen, A., Laureys, S., & De Tiège, X. (2008).

[Functional neuroimaging (fMRI, PET and MEG): What do we measure?]. *Revue Medicale De Liege*, *63*(5–6), 231–237.

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnins.2013.00267

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLOS Biology*, *11*(12), e1001752. https://doi.org/10.1371/journal.pbio.1001752

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, *98*(2), 694–699. https://doi.org/10.1073/pnas.98.2.694

Guan, A., Wang, S., Huang, A., Qiu, C., Li, Y., Li, X., Wang, J., Wang, Q., & Deng, B. (2022). The role of gamma oscillations in central nervous system diseases: Mechanism and treatment. *Frontiers in Cellular Neuroscience*, *16*. https://www.frontiersin.org/articles/10.3389/fncel.2022.962957

Halgren, M., Ulbert, I., Bastuji, H., Fabó, D., Erőss, L., Rey, M., Devinsky, O., Doyle, W. K., Mak-McCully, R., Halgren, E., Wittner, L., Chauvel, P., Heit, G., Eskandar, E., Mandell, A., & Cash, S. S. (2019). The generation and propagation of the human alpha rhythm. *Proceedings of the National Academy of Sciences*, *116*(47), 23772–23782. https://doi.org/10.1073/pnas.1913092116

Hari, R. (2011). Magnetoencephalography: Methods and applications. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. https://research.aalto.fi/en/publications/magnetoencephalography-methods-and-applications

Hari, R., & Puce, A. (2017). *MEG–EEG Primer*. Oxford University Press.

> https://doi.org/10.1093/med/9780190497774.003.0001

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010).

> Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the*
> *National Academy of Sciences*, *107*(38), 16494–16499.
> https://doi.org/10.1073/pnas.1008121107

He, B. J. (2014). Scale-free brain activity: Past, present, and future. *Trends in Cognitive Sciences*,

> *18*(9), 480–487. https://doi.org/10.1016/j.tics.2014.04.003

He, B. J., Zempel, J. M., Snyder, A. Z., & Raichle, M. E. (2010). The Temporal Structures and

> Functional Significance of Scale-free Brain Activity. *Neuron*, *66*(3), 353–369.
> https://doi.org/10.1016/j.neuron.2010.04.020

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335).

> Wiley.

Helfrich, R. F., Lendner, J. D., & Knight, R. T. (2021). Aperiodic sleep networks promote memory

> consolidation. *Trends in Cognitive Sciences*, *25*(8), 648–659.
> https://doi.org/10.1016/j.tics.2021.04.009

Hermes, D., Miller, K. J., Wandell, B. A., & Winawer, J. (2015). Stimulus Dependence of Gamma

> Oscillations in Human Visual Cortex. *Cerebral Cortex (New York, NY)*, *25*(9), 2951–2959.
> https://doi.org/10.1093/cercor/bhu091

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception.

> *Trends in Cognitive Sciences*, *4*(4), 131–138. https://doi.org/10.1016/S1364-6613(00)01463-
> 7

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews.*

> *Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception & Psychophysics*, *72*(5), 1218–1227. https://doi.org/10.3758/APP.72.5.1218

Hooker, C. I., & Knight, R. T. (2006). The role of lateral orbitofrontal cortex in the inhibitory control of emotion. In D. Zald & S. Rauch (Eds.), *The Orbitofrontal Cortex* (1st ed., pp. 307–324). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780198565741.003.0012

Ibarra Chaoul, A., & Siegel, M. (2021). Cortical correlation structure of aperiodic neuronal population activity. *Neuroimage*, *245*, 118672. https://doi.org/10.1016/j.neuroimage.2021.118672

Jensen, O., Spaak, E., & Zumer, J. M. (2014). Human brain oscillations: From physiological mechanisms to analysis and cognition. In *Magnetoencephalography* (Vol. 9783642330452, pp. 359–403). Springer. https://doi.org/10.1007/978-3-642-33045-2_17

Kawakatsu, M. (2003). Application of ICA to MEG noise reduction.

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biology*, *16*(3), e2004473. https://doi.org/10.1371/journal.pbio.2004473

Khanna, P., & Carmena, J. M. (2017). Beta band oscillations in motor cortex reflect neural population signals that delay movement onset. *eLife*, *6*, e24573. https://doi.org/10.7554/eLife.24573

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Lange, J., Keil, J., Schnitzler, A., van Dijk, H., & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behavioural Brain Research*, *271*, 294–301. https://doi.org/10.1016/j.bbr.2014.06.015

Larson, E., & Taulu, S. (2018). Reducing Sensor Noise in MEG and EEG Recordings Using Oversampled Temporal Projection. *IEEE Transactions on Bio-Medical Engineering*, *65*(5), 1002–1013. https://doi.org/10.1109/TBME.2017.2734641

Leonard, M. K., & Chang, E. F. (2014). Dynamic speech representations in the human temporal lobe. *Trends in Cognitive Sciences*, *18*(9), 472–479. https://doi.org/10.1016/j.tics.2014.05.001

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. https://doi.org/10.1037/h0020279

López-Barroso, D., & de Diego-Balaguer, R. (2017). Language Learning Variability within the Dorsal and Ventral Streams as a Cue for Compensatory Mechanisms in Aphasia Recovery. *Frontiers in Human Neuroscience*, *11*. https://www.frontiersin.org/articles/10.3389/fnhum.2017.00476

Lowet, E., Roberts, M. J., Bonizzi, P., Karel, J., & Weerd, P. D. (2016). Quantifying Neural Oscillatory Synchronization: A Comparison between Spectral Coherence and Phase-Locking Value Approaches. *PLOS ONE*, *11*(1), e0146443. https://doi.org/10.1371/journal.pone.0146443

Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., & Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-017-02791-8

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Mably, A. J., & Colgin, L. L. (2018). Gamma oscillations in cognitive disorders. *Current Opinion in Neurobiology*, *52*, 182–187. https://doi.org/10.1016/j.conb.2018.07.009

Madell, J. R., Hewitt, J. G., & Rotfleisch, S. (2018). Auditory Perception vs. Speech Production in Children. *The Hearing Journal*, *71*(12), 14. https://doi.org/10.1097/01.HJ.0000550393.51777.97

Mäkelä, S., Kujala, J., & Salmelin, R. (2022). Removing ocular artifacts from magnetoencephalographic data on naturalistic reading of continuous texts. *Frontiers in Neuroscience*, *16*, 974162. https://doi.org/10.3389/fnins.2022.974162

Maniscalco, B., Lee, J. L., Abry, P., Lin, A., Holroyd, T., & He, B. J. (2018). Neural Integration of Stimulus History Underlies Prediction for Naturalistically Evolving Sequences. *Journal of Neuroscience*, *38*(6), 1541–1557. https://doi.org/10.1523/JNEUROSCI.1779-17.2017

Manning, L., & Thomas-Antérion, C. (2011). Marc Dax and the discovery of the lateralisation of language in the left cerebral hemisphere. *Revue Neurologique*, *167*(12), 868–872. https://doi.org/10.1016/j.neurol.2010.10.017

Mantini, D., Franciotti, R., Romani, G. L., & Pizzella, V. (2008). Improving MEG source localizations: An automated method for complete artifact removal based on independent component analysis. *NeuroImage*, *40*(1), 160–173. https://doi.org/10.1016/j.neuroimage.2007.11.022

Martini, E. (2022, July 25). *Unbiased two-windows approach for Welch's method*. https://www.semanticscholar.org/paper/2f70148cc427806f0fb3504b27df0a67041af78f

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The Cortical Representation of Speech.

Journal of Cognitive Neuroscience, *5*(4), 467–479.

https://doi.org/10.1162/jocn.1993.5.4.467

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of

premotor cortex in speech perception. *Current Biology: CB*, *17*(19), 1692–1696.

https://doi.org/10.1016/j.cub.2007.08.064

Miller, K. J., Honey, C. J., Hermes, D., Rao, R. P., denNijs, M., & Ojemann, J. G. (2014).

Broadband changes in the cortical surface potential track activation of functionally diverse

neuronal populations. *NeuroImage*, *85*(0 2), 711–720.

https://doi.org/10.1016/j.neuroimage.2013.08.070

Miller, K. J., Sorensen, L. B., Ojemann, J. G., & Nijs, M. den. (2009). Power-Law Scaling in the

Brain Surface Electric Potential. *PLOS Computational Biology*, *5*(12), e1000609.

https://doi.org/10.1371/journal.pcbi.1000609

Mohseni, H. R., Woolrich, M. W., Kringelbach, M. L., Luckhoo, H., Smith, P., & Aziz, T. Z.

(2011). *Fusion of Magnetometer and Gradiometer Sensors of MEG in a Beamforming

Framework*.

Molina, J. L., Voytek, B., Thomas, M. L., Joshi, Y. B., Bhakta, S. G., Talledo, J. A., Swerdlow, N.

R., & Light, G. A. (2020). Memantine Effects on Electroencephalographic Measures of

Putative Excitatory/Inhibitory Balance in Schizophrenia. *Biological Psychiatry: Cognitive

Neuroscience and Neuroimaging*, *5*(6), 562–568. https://doi.org/10.1016/j.bpsc.2020.02.004

Muthukumaraswamy, S. D., & Singh, K. D. (2013). Visual gamma oscillations: The effects of

stimulus type, visual field coverage and stimulus motion on MEG and EEG recordings.

NeuroImage, *69*, 223–230. https://doi.org/10.1016/j.neuroimage.2012.12.038

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: A metaphor? *Philosophical Transactions of the

Royal Society B: Biological Sciences*, *369*(1658), 20130396.

https://doi.org/10.1098/rstb.2013.0396

Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex*, *22*, 2466–2477. https://doi.org/10.1093/cercor/bhr325

Ostlund, B. D., Alperin, B. R., Drew, T., & Karalunas, S. L. (2021). Behavioral and cognitive correlates of the aperiodic (1/f-like) exponent of the EEG power spectrum in adolescents with and without ADHD. *Developmental Cognitive Neuroscience*, *48*, 100931. https://doi.org/10.1016/j.dcn.2021.100931

Ou, J., & Law, S.-P. (2019). Top-down and bottom-up mechanisms as reflected by beta and gamma oscillations in speech perception: An individual-difference approach. *Brain and Language*, *199*, 104700. https://doi.org/10.1016/j.bandl.2019.104700

Pantev, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., & Gallen, C. (1991). Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(20), 8996–9000.

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, NY)*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Peterson, E., Rosen, B., Campbell, A., Belger, A., & Voytek, B. (2018). *1/f neural noise is a better predictor of schizophrenia than neural oscillations*. https://doi.org/10.1101/113449

Piastra, M. C., Nüßing, A., Vorwerk, J., Clerc, M., Engwer, C., & Wolters, C. H. (2021). A comprehensive study on electroencephalography and magnetoencephalography sensitivity to cortical and subcortical sources. *Human Brain Mapping*, *42*(4), 978–992. https://doi.org/10.1002/hbm.25272

Picton, T. W., & Picton, T. W. (2010). *Human Auditory Evoked Potentials*. Plural Publishing, Inc. http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=1886781

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral

lateralization as 'asymmetric sampling in time.' *Speech Communication*, *41*(1), 245–255.

https://doi.org/10.1016/S0167-6393(02)00107-3

Poeppel, D. (2015). Speech Perception. In A. W. Toga (Ed.), *Brain Mapping* (pp. 429–434).

Academic Press. https://doi.org/10.1016/B978-0-12-397025-1.00264-5

Poeppel, D., Assaneo, M.F. (2020). Speech rhythms and their neural foundations. *Nat Rev Neurosci*

21, 322–334 (2020). https://doi.org/10.1038/s41583-020-0304-4

Rahimpour, S., Haglund, M. M., Friedman, A. H., & Duffau, H. (2019). History of awake mapping

and speech and language localization: From modules to networks. *Neurosurgical Focus*,

47(3), E4. https://doi.org/10.3171/2019.7.FOCUS19347

Roberts, B., Summers, R. J., & Bailey, P. J. (2011). The intelligibility of noise-vocoded speech:

Spectral information available from across-channel comparison of amplitude envelopes.

Proceedings of the Royal Society B: Biological Sciences, *278*(1711), 1595–1600.

https://doi.org/10.1098/rspb.2010.1554

Samaha, J., & Postle, B. R. (2015). The Speed of Alpha-Band Oscillations Predicts the Temporal

Resolution of Visual Perception. *Current Biology*, *25*(22), 2985–2990.

https://doi.org/10.1016/j.cub.2015.10.007

Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E. G., & Belin, P. (2015). Dorsal

and Ventral Pathways for Prosody. *Current Biology*, *25*(23), 3079–3085.

https://doi.org/10.1016/j.cub.2015.10.009

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R.,

Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008).

Ventral and dorsal pathways for language. *Proceedings of the National Academy of*

Sciences, *105*(46), 18035–18040. https://doi.org/10.1073/pnas.0805234105

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006

Schomers, M. R., & Pulvermüller, F. (2016). Is the Sensorimotor Cortex Relevant for Speech Perception and Understanding? An Integrative Review. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00435

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science (New York, N.Y.)*, *270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Shtyrov, Y., & Pulvermüller, F. (2007). Early MEG Activation Dynamics in the Left Temporal and Inferior Frontal Cortex Reflect Semantic Context Integration. *Journal of Cognitive Neuroscience*, *19*(10), 1633–1642. https://doi.org/10.1162/jocn.2007.19.10.1633

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, *25*(1), 76–89. https://doi.org/10.1016/j.neuroimage.2004.11.006

Snipes, S., Krugliakova, E., Meier, E., & Huber, R. (2022). The Theta Paradox: 4-8 Hz EEG Oscillations Reflect Both Sleep Pressure and Cognitive Control. *Journal of Neuroscience*, *42*(45), 8569–8586. https://doi.org/10.1523/JNEUROSCI.1063-22.2022

Specht, K. (2014). Neuronal basis of speech comprehension. *Hearing Research*, *307*, 121–135. https://doi.org/10.1016/j.heares.2013.09.011

Stone, J. V. (2002). *Independent component analysis: An introduction*.

Sun, L., Hämäläinen, M. S., & Okada, Y. (2018). Noise cancellation for a whole-head magnetometer-based MEG system in hospital environment. *Biomedical Physics & Engineering Express*, *4*(5). https://doi.org/10.1088/2057-1976/aad627

Swettenham, J. B., Muthukumaraswamy, S. D., & Singh, K. D. (2009). Spectral properties of induced and evoked gamma oscillations in human early visual cortex to moving and stationary stimuli. *Journal of Neurophysiology*, *102*(2), 1241–1253. https://doi.org/10.1152/jn.91044.2008

Tesche, C. D., & Karhu, J. (2000). Theta Oscillations Index Human Hippocampal Activation during a Working Memory Task. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(2), 919–924.

Thomas, B., & John, S.K. (2015). Power Spectral Density Computation using Modified Welch Method.

Treacher, A. H., Garg, P., Davenport, E., Godwin, R., Proskovec, A., Bezerra, L. G., Murugesan, G., Wagner, B., Whitlow, C. T., Stitzel, J. D., Maldjian, J. A., & Montillo, A. A. (2021). MEGnet: Automatic ICA-based artifact removal for MEG using spatiotemporal convolutional neural networks. *NeuroImage*, *241*, 118402. https://doi.org/10.1016/j.neuroimage.2021.118402

van der Burght, C. L., Numssen, O., Schlaak, B., Goucha, T., & Hartwigsen, G. (2023). Differential contributions of inferior frontal gyrus subregions to sentence processing guided by intonation. *Human Brain Mapping*, *44*(2), 585–598. https://doi.org/10.1002/hbm.26086

van Heumen, S., Moreau, J. T., Simard-Tremblay, E., Albrecht, S., Dudley, R. WR., & Baillet, S. (2021). Case Report: Aperiodic Fluctuations of Neural Activity in the Ictal MEG of a Child With Drug-Resistant Fronto-Temporal Epilepsy. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnhum.2021.646426

Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *The Journal of the Acoustical Society of America*, *142*(4), 1976–1989. https://doi.org/10.1121/1.5006179

Varvatsoulias, G. (2013). The Physiological Processes Underpinning PET and fMRI Techniques With an Emphasis on the Temporal and Spatial Resolution of These Methods. *Psychological Thought*, *6*, 173–195. https://doi.org/10.5964/psyct.v6i2.75

Voss, R. F., & Clarke, J. (1975). '1/fnoise' in music and speech. *Nature*, *258*(5533), Article 5533. https://doi.org/10.1038/258317a0

Voytek, B., & Knight, R. T. (2015). Dynamic network communication as a unifying neural basis for cognition, development, aging, and disease. *Biological Psychiatry*, *77*(12), 1089–1097. https://doi.org/10.1016/j.biopsych.2015.04.016

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1/f Neural Electrophysiological Noise. *The Journal of Neuroscience*, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989–994. https://doi.org/10.1016/S0028-3932(02)00316-0

Westad, F., & Kermit, M. (2009). 2.14—Independent Component Analysis. In S. D. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive Chemometrics* (pp. 227–248). Elsevier. https://doi.org/10.1016/B978-044452701-1.00045-4

Wilson, L. E., Castanheira, J. da S., & Baillet, S. (2022). *Time-resolved parameterization of aperiodic and periodic brain activity* (p. 2022.01.21.477243). bioRxiv. https://doi.org/10.1101/2022.01.21.477243

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), Article 7. https://doi.org/10.1038/nn1263

Zhou, Y. J., Iemi, L., Schoffelen, J.-M., Lange, F. P. de, & Haegens, S. (2021). Alpha Oscillations Shape Sensory Representation and Perceptual Sensitivity. *Journal of Neuroscience*, *41*(46), 9581–9592. https://doi.org/10.1523/JNEUROSCI.1114-21.2021

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., & Schroeder, C. E. (2013). Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a 'Cocktail Party.' *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

APPENDIXES

Appendix A Audio samples used

Audio clip	Difficulty (1=easy, 2=difficult)	source
Helppo_maailma_2	1	Helppo_maailma_2
Helppo_karhu_1	1	Helppo_karhu_1
Helppo_karhu_2	1	Helppo_karhu_2
Helppo_robotti_1	1	Helppo_robotti_1
Helppo_robotti_2	1	Helppo_robotti_2
Vaikea_perhonen_1	2	Vaikea_perhonen_1
Vaikea_perhonen_2	2	vaikea_perhonen_2
Vaikea_perhonen_4	2	Vaikea_perhonen_4
Vaikea_maailma_3	2	Vaikea_maailma_3
Vaikea_maailma_4	2	Vaikea_maailma_4

Appendix B Power values (tesla²/m²) for brain areas where statistically significant differences between speech settings were found
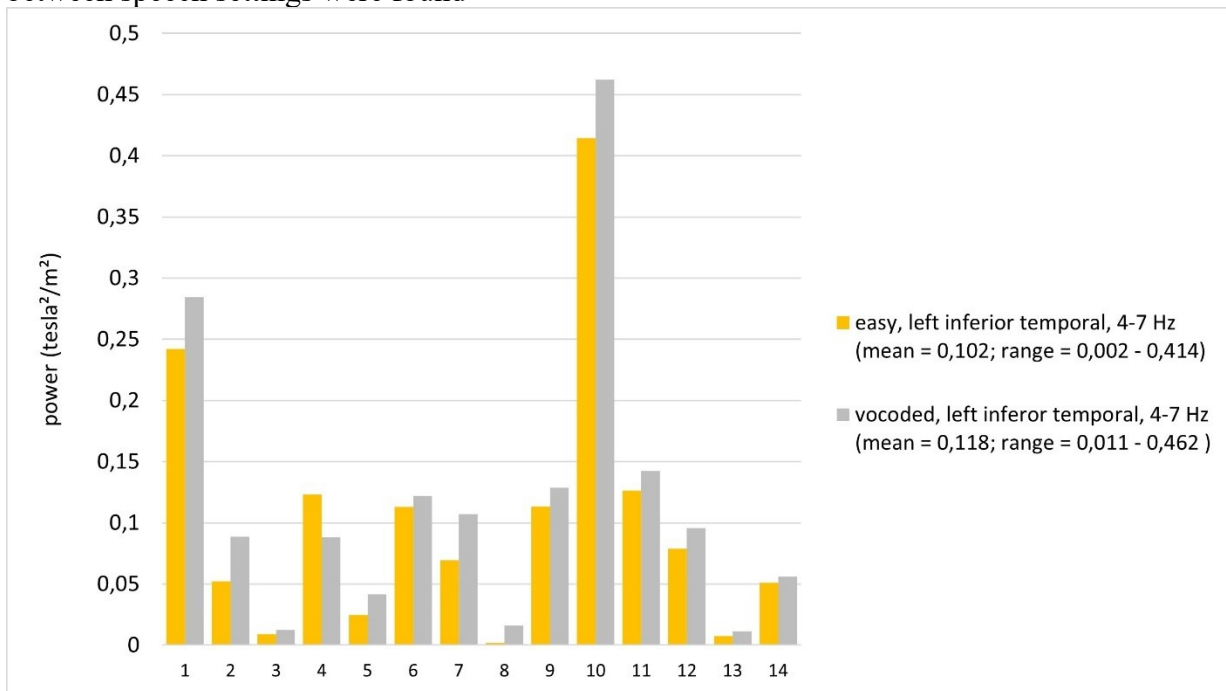
B1. Power values of theta oscillations (4 – 7 Hz), easy and vocoded speech, (4-7 Hz), left inferior temporal region.
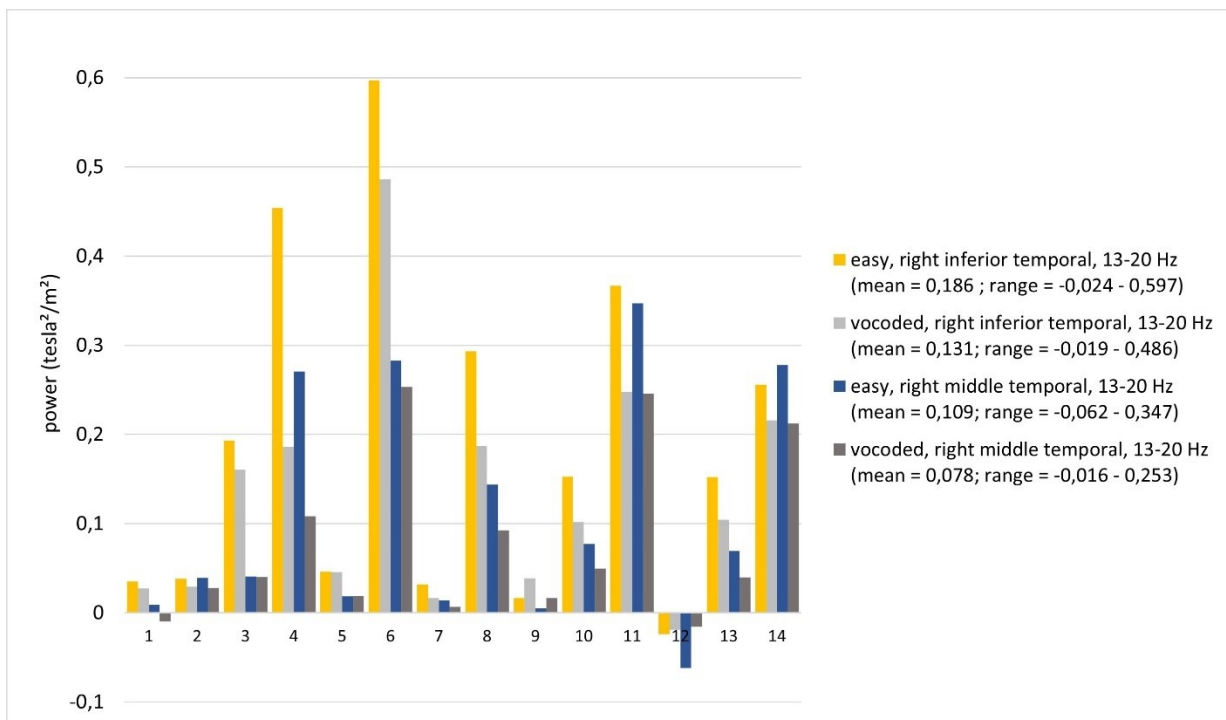
B2. Power values of beta oscillations (13 – 20 Hz), easy and vocoded speech, right inferior and middle temporal regions.
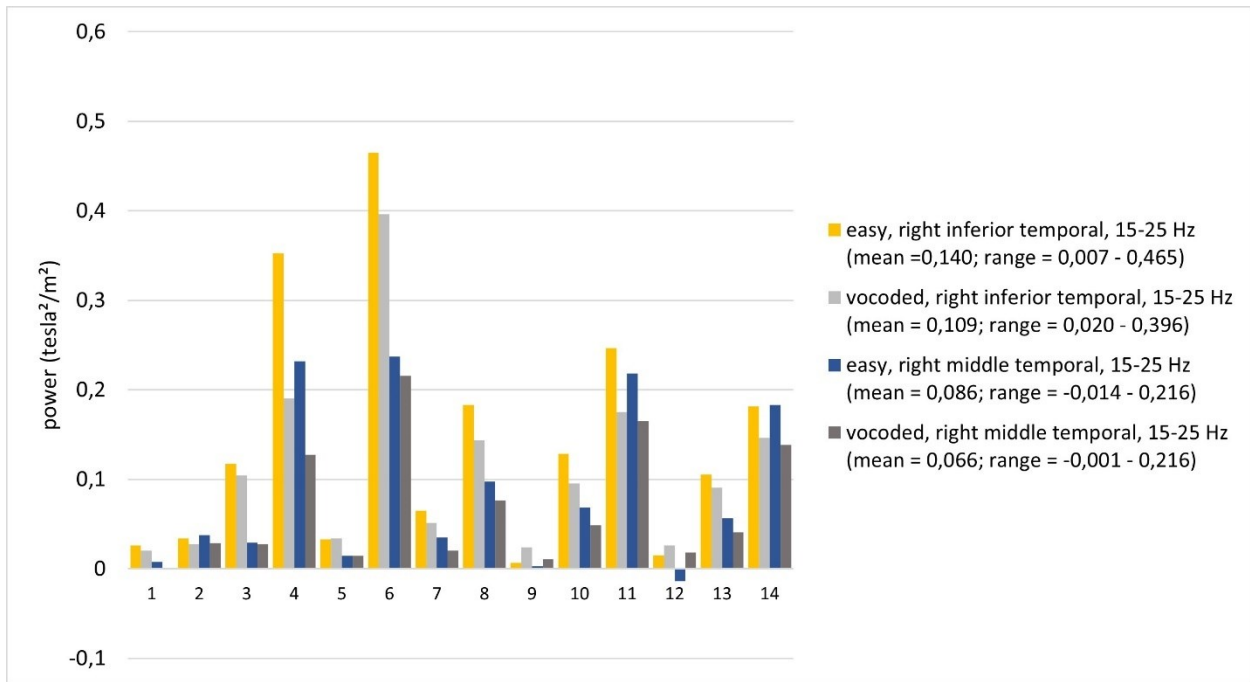
B3. Power values of beta oscillations (15-25 Hz), easy and vocoded speech, right middle and inferior temporal regions.
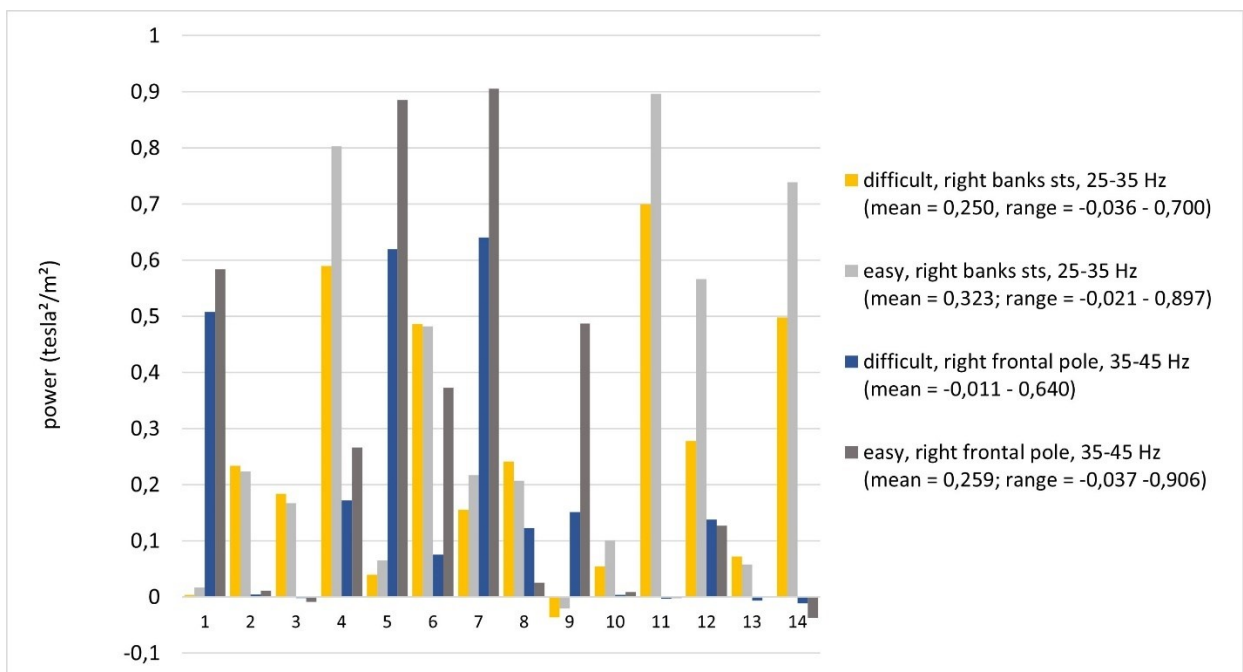
B4. Power values of gamma oscillations (25-35 Hz, 35-45 Hz), difficult and easy speech, right banks sts and frontal pole
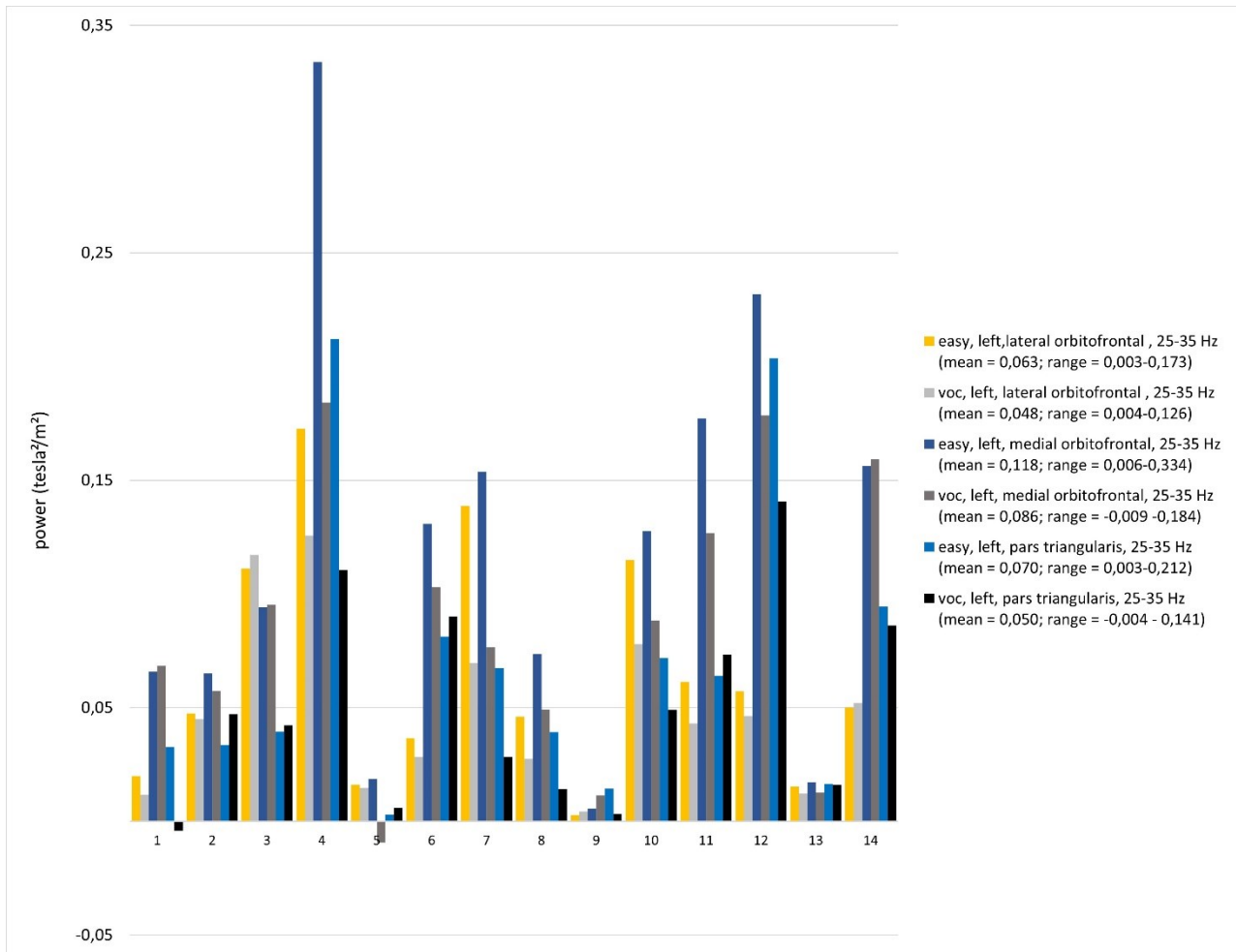
B5. Power values of gamma oscillations (25-35 Hz), easy and vocoded speech, left lateral and medial orbitofrontal and pars triangularis.
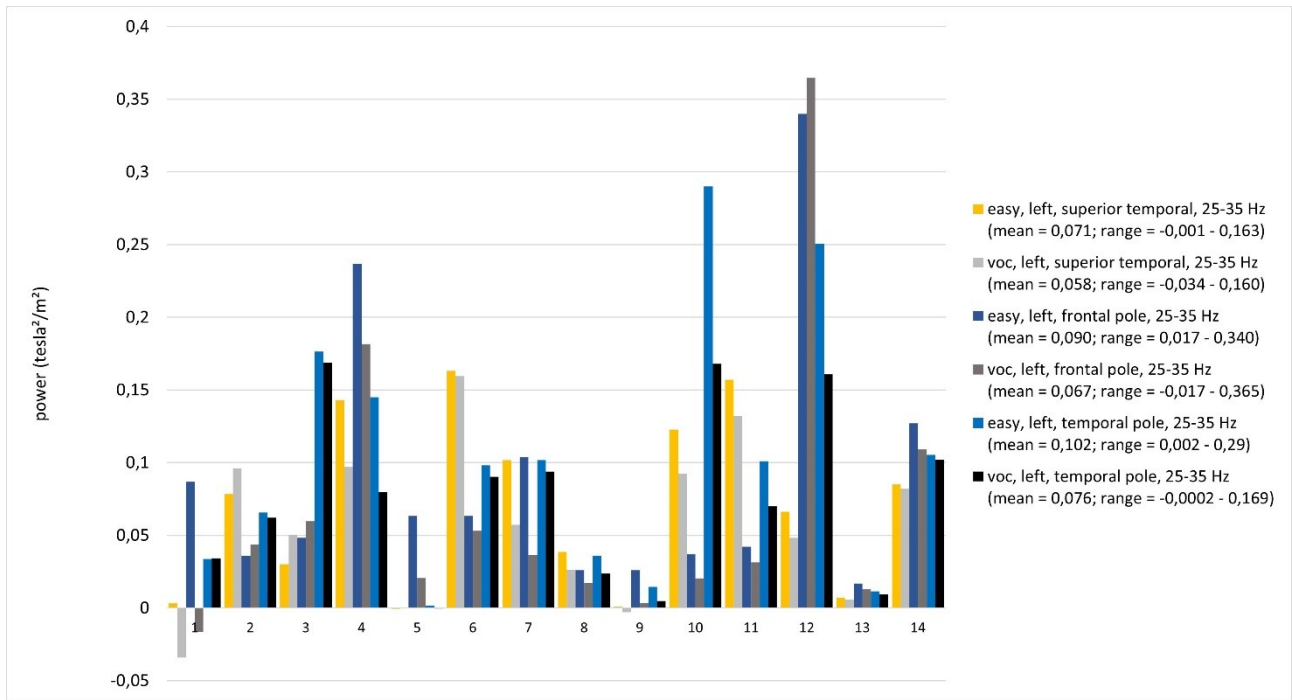
B6. Power values of gamma oscillations (25-35 Hz), easy and vocoded speech, left superior temporal, frontal pole, and temporal pole
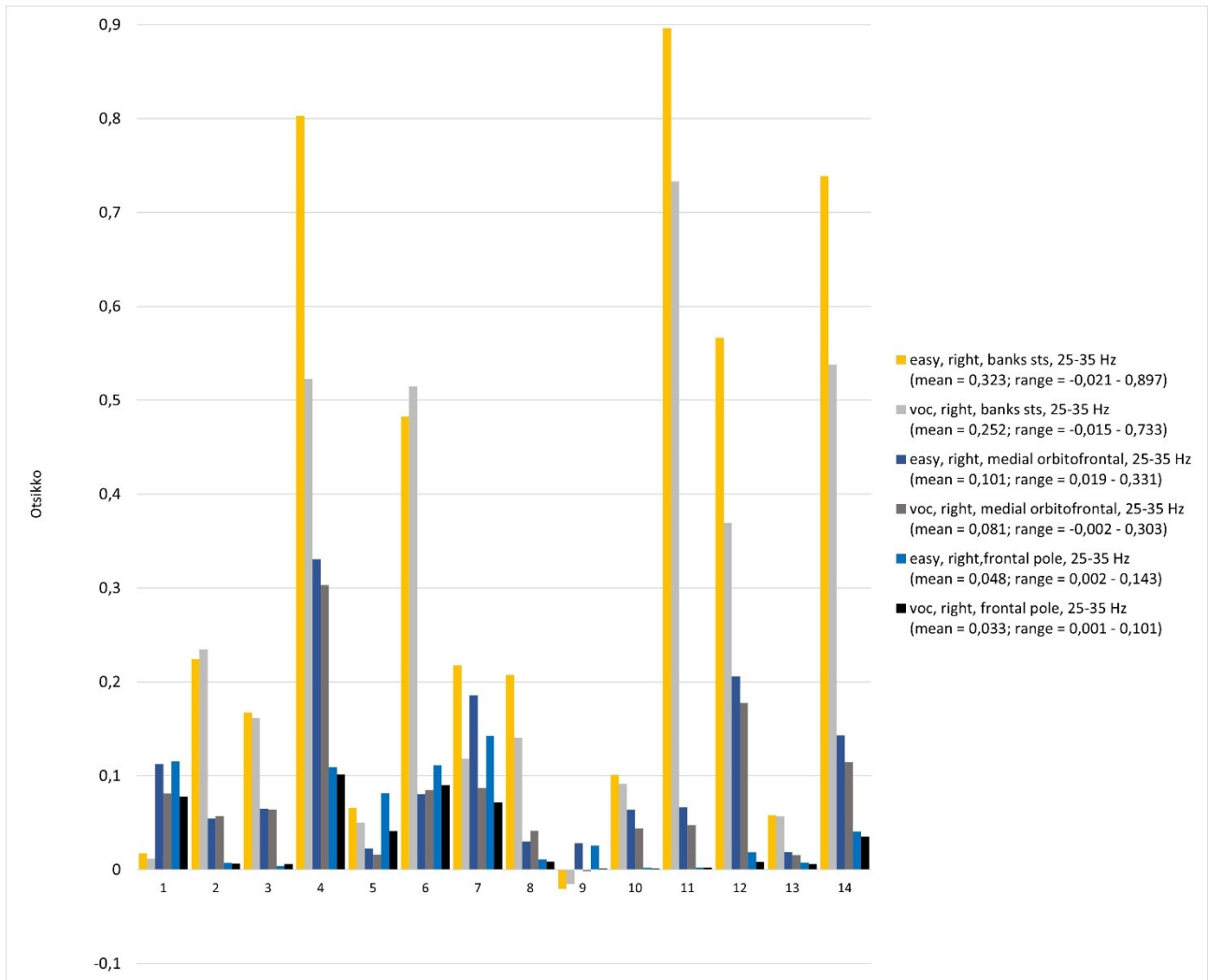
B7. Power values of gamma oscillations (25-35 Hz), easy and vocoded speech, right banks sts, medial orbitofrontal, and frontal pole
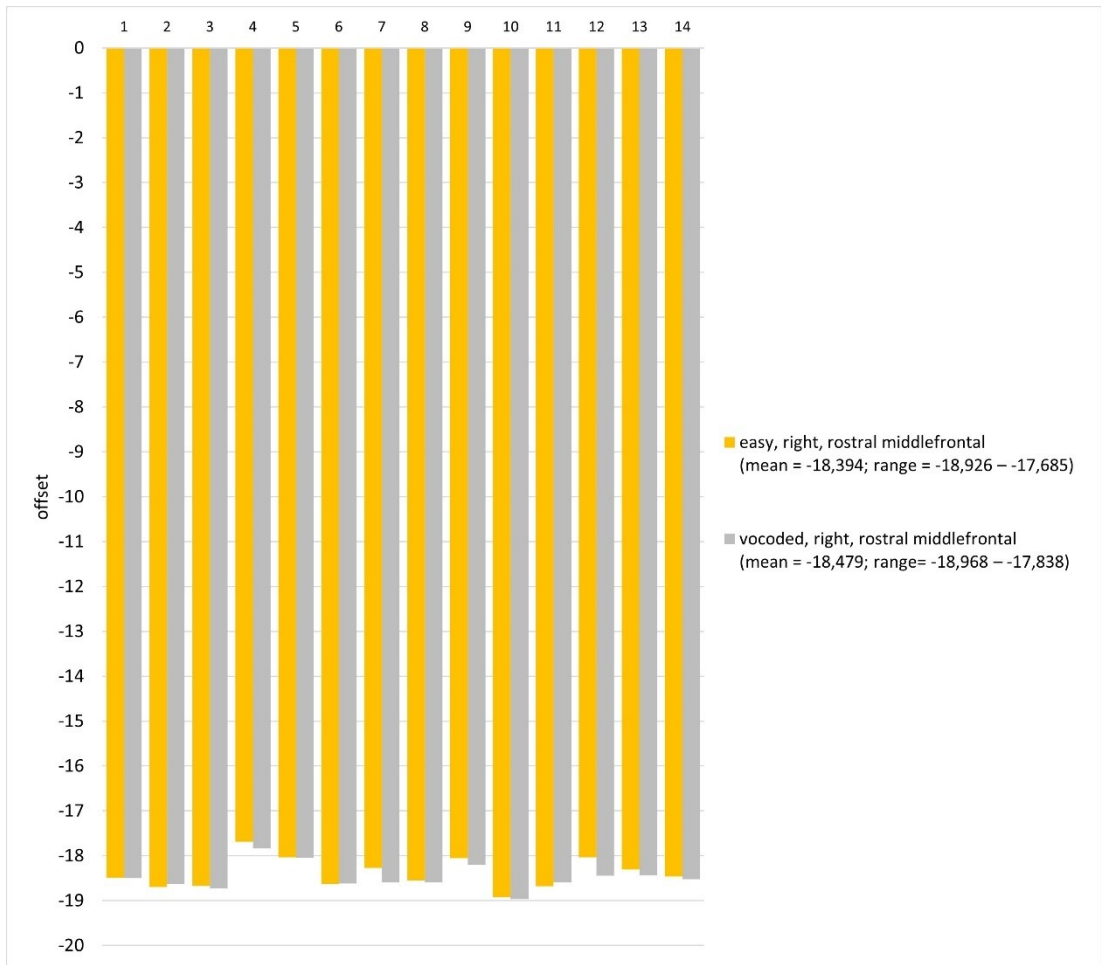
Appendix C. Exponent and offset values for brain areas where statistically significant differences between speech conditions were found
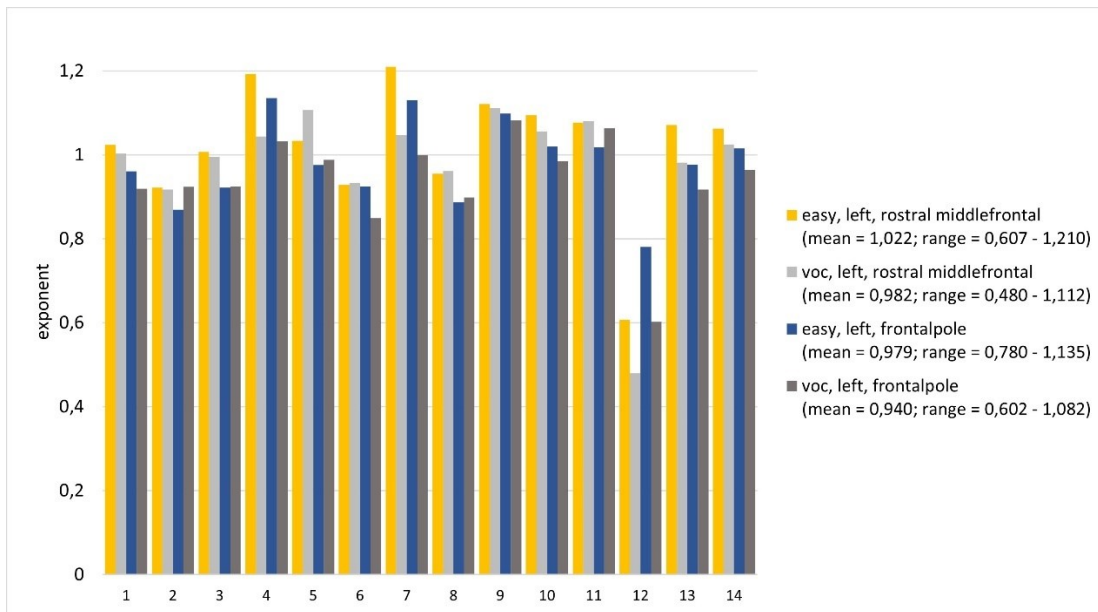
*C1 Exponent values, easy and vocoded, left rostral middle frontal and frontal pole*
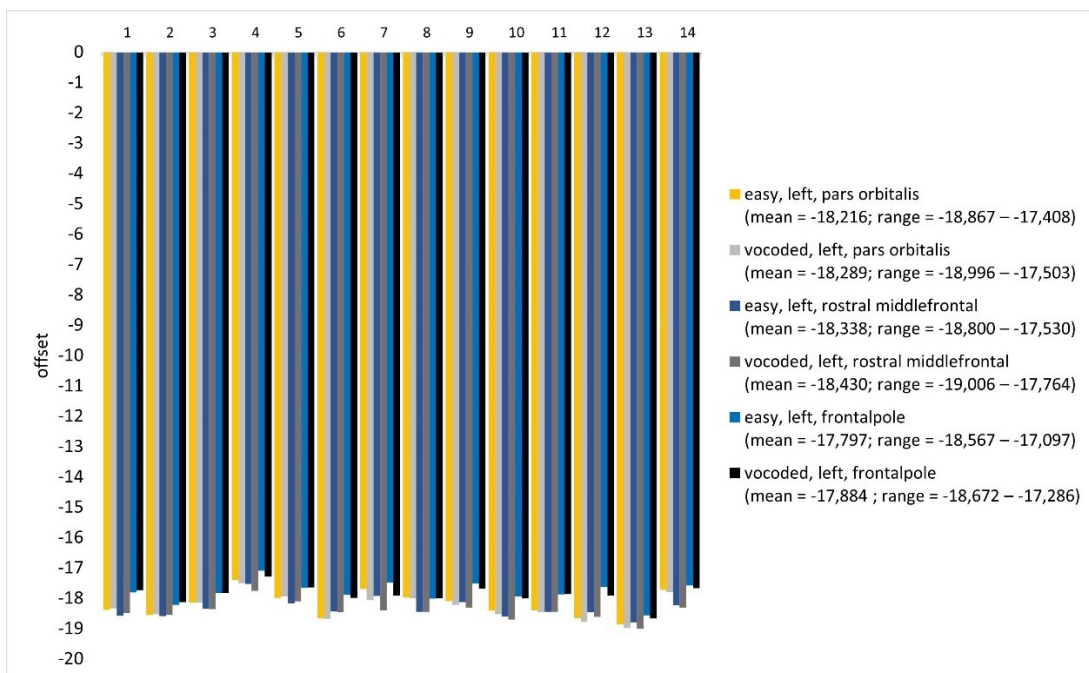
C2 Offset values, easy and vocoded, left pars orbitalis, rostral middlefrontal and frontalpole regions
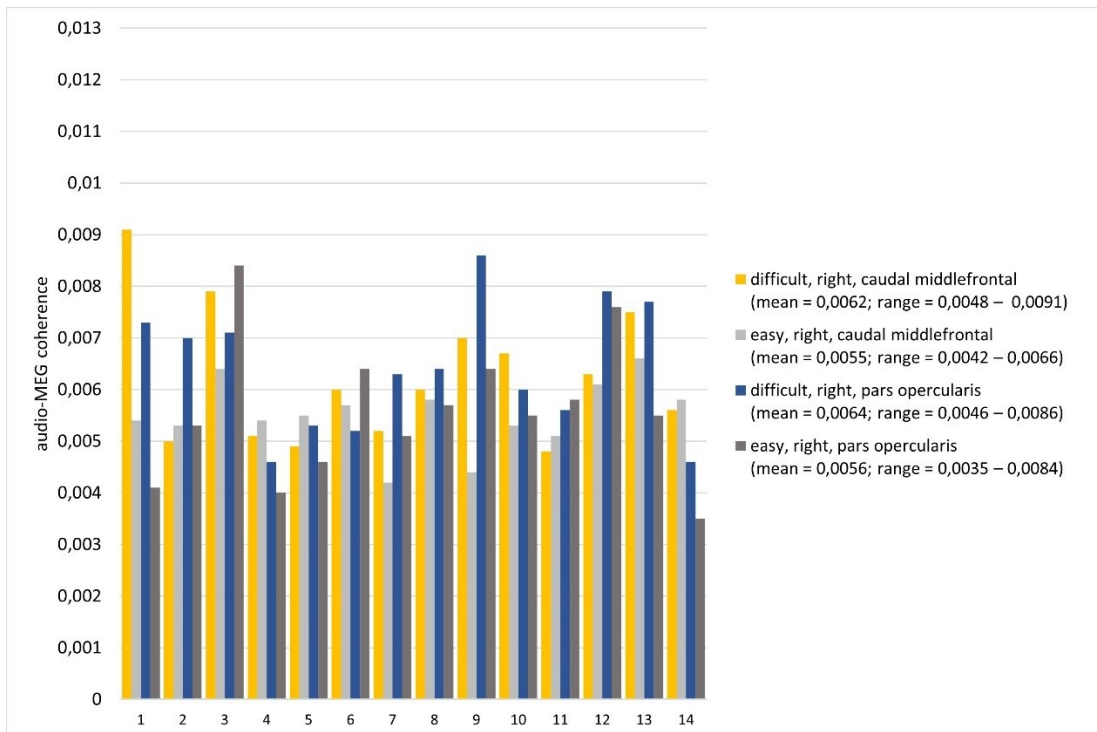
C3 Offset values, easy and vocoded, right, rostral middlefrontal regions

Appendix D Audio-MEG coherence values for brain areas where statistically significant differences between speech conditions were found
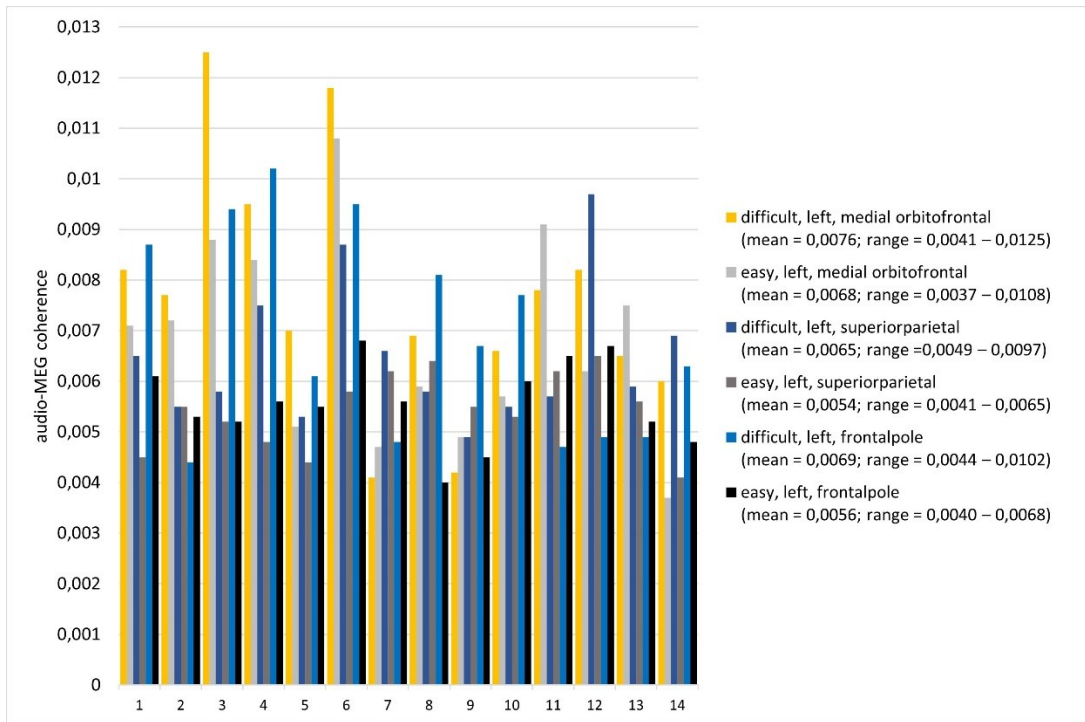
D1 Audio-MEG coherence values, difficult and easy speech, left, medial orbitofrontal, superiorparietal and frontalpole regions

D2 Audio-MEG coherence values, difficult and easy speech, right, caudal middlefrontal and pars opercularis regions
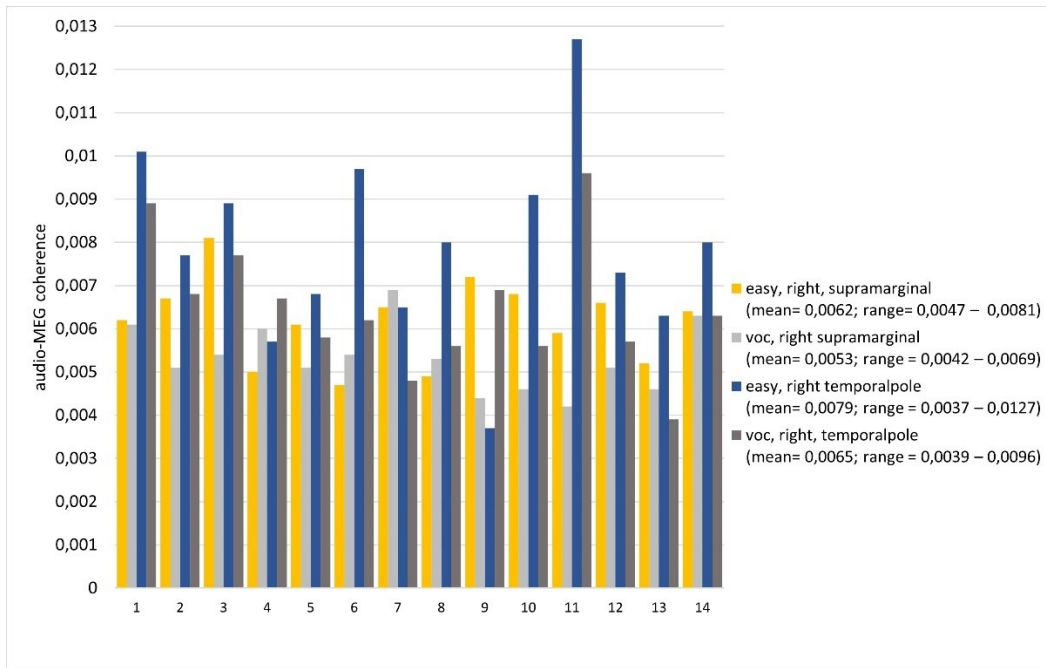
D3 Audio-MEG coherence values, easy and vocoded speech, right supramarginal and temporalpole regions