






Alpha Oscillations During Effortful Continuous Speech: From Scalp EEG to Ear-EEG

Tirdad Seifi Ala , Emina Alickovic , Alvaro Fuentes Cabrera , William M. Whitmer , Lauren V. Hadley, Mike L. Rank , *Member, IEEE*, Thomas Lunner, and Carina Graversen

Abstract—Objective: The purpose of this study was to investigate alpha power as an objective measure of effortful listening in continuous speech with scalp and ear-EEG. **Methods:** Scalp and ear-EEG were recorded simultaneously during presentation of a 33-s news clip in the presence of 16-talker babble noise. Four different signal-to-noise ratios (SNRs) were used to manipulate task demand. The effects of changes in SNR were investigated on alpha event-related synchronization (ERS) and desynchronization (ERD). Alpha activity was extracted from scalp EEG using different referencing methods (common average and symmetrical bipolar) in different regions of the brain (parietal and temporal) and ear-EEG. **Results:** Alpha ERS decreased with decreasing SNR (i.e., increasing task demand) in both scalp and ear-EEG. Alpha ERS was also positively correlated to behavioural performance which was based on the questions regarding the contents of the speech. **Conclusion:** Alpha ERS/ERD is better suited to track performance of a continuous speech than listening effort. **Significance:** EEG alpha power in continuous speech may indicate of how well the speech was perceived and it can be measured with both scalp and Ear-EEG.

Index Terms—Alpha power, EEG, ear-EEG, Listening Effort.

I. INTRODUCTION

LISTENING effort is referred as “the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a [listening] task” [1]. Listening effort may lead to unwanted consequences, such as fatigue or difficulty in comprehension and remembering the speech [2]. Physiological measurements including pupillometry [3], [4], [5] or electroencephalography (EEG) [6], [7] have shown an inverted U-shaped pattern of objective listening effort as a function of task demand. This inverted U-shape demonstrates that increased difficulty of a task leads to increased effort as long as there are enough cognitive resources available to benefit performance of the task. If the task becomes too difficult, then the listener will disengage from the listening task, and that leads to a drop in the objective measure of effort [1].

EEG has been widely used in the literature to measure effort objectively, whether in auditory [7], [8], [9], [10], [11] or non-auditory [12], [13], [14] studies. Specifically, alpha oscillations have been the key measure in most of these studies concerning effort. How alpha power plays a role in the brain during an effortful task, however, is not completely understood. It has been suggested that an increase in alpha oscillations inhibits task-irrelevant areas [15] and can be used as a neural correlate of listening effort (see review [16]). The problem with alpha power is that it is not merely associated with effort. Alpha power also correlates with optimal task performance [17] or it might be used to predict speech intelligibility [18]. Given that effort and performance could vary independently of each other [19], [20], [21] this raises an important question of when alpha power reflects both or either of these concepts.

Traditional EEG, which is also known as scalp EEG, has enjoyed wide attention from brain researchers, due to its non-invasiveness and excellent temporal resolution, along with several other advantages [22]. However, despite high quality recordings by scalp EEG, it is not yet suitable as a wearable technology, instead restricting the wearer’s movement and limiting measurement to laboratory experiments [23]. However, EEG signals can now be picked up by electrodes placed inside the ears, which is known as ear-EEG, allowing ambulatory measurement of brain activity both in the laboratory and at home [23], [24]. Despite the attraction of ear-EEG as a wearable technology, it still suffers from poorer spatial resolution compared to scalp EEG, as the signals are limited to the cortical areas close to the ears [25]. Nonetheless, researchers have shown that traditional analyses,

Manuscript received 23 March 2022; revised 17 August 2022; accepted 3 October 2022. Date of publication 13 October 2022; date of current version 21 March 2023. The work of Tirdad Seifi Ala was supported by the European Union’s Horizon 2020 Research and Innovation Programme under Marie Skłodowska-Curie Grant agreement No 765329. The work of William M. Whitmer and Lauren V. Hadley was supported in part by the Medical Research Council under Grants MR/S003576/1 and MR/X003620/1, and in part by the Chief Scientist Office of the Scottish Government. The work of Lauren V. Hadley was also supported by the UKRI Future Leaders Fellowship under Grant MR/T041471/1. (Corresponding author: Tirdad Seifi Ala.)

Tirdad Seifi Ala is with the Eriksholm Research Centre, Oticon A/S, Denmark and with the Hearing Sciences – Scottish Section, Division of Mental Health and Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham NG7 2RD, U.K. (e-mail: tial@eriksholm.com).

Emina Alickovic is with the Eriksholm Research Centre, Oticon A/S, Denmark and with the Department of Electrical Engineering, Linköping University, Sweden.

Alvaro Fuentes Cabrera and Mike L. Rank are with the T&W Engineering A/S, Denmark.

William M. Whitmer and Lauren V. Hadley are with the Hearing Sciences – Scottish Section, Division of Mental Health and Clinical Neuroscience, School of Medicine, University of Nottingham, U.K.

Thomas Lunner and Carina Graversen were with Eriksholm Research Centre, Oticon A/S, Denmark.

Digital Object Identifier 10.1109/TBME.2022.3214428

such as P300 detection [24], frequency-domain representation [26], steady-state evoked potential [27], or decoding selective attention [28], [29], [30], commonly studied via scalp EEG, are also evident in ear-EEG. However, listening effort has not yet been explored using ear-EEG.

In the present study, we investigated the effects of changing task demand by manipulating the signal-to-noise ratio (SNR) of a continuous speech-in-noise task on EEG data. We used continuous speech as it represents a more realistic listening scenario than traditional short-sentence paradigms. Both scalp and ear-EEG were recorded simultaneously during this task. We previously conducted two experiments with similar long, continuous speech to investigate the effects of varying SNR on alpha power (see [31], [32]), using only scalp EEG. In both studies we observed that more demanding conditions led to decreased alpha power. In this study, we aimed to use a wider range of SNR compared to the two aforementioned studies to capture a bigger picture of alpha changes across different task demands. Therefore, we implemented four different SNRs (-16 , -8 , -4 , $+8$ dB), creating very low to very high demand conditions for a normal-hearing person listening to continuous speech. Based on the results of our two previous experiments, we expected that alpha power would increase with increasing SNR in scalp EEG. Using ear-EEG for the first time in this paradigm, we then investigated how the measured alpha power in scalp EEG and ear-EEG compared to each other.

II. METHODS

A. Participants

Sixteen (4 females) normal-hearing Danish-speaking adults (average age of 42.4 ± 11.4 years) participated in this study. One male participant was excluded due to a problem in sending triggers to the EEG device. All participants signed a written consent form before the experiment. Ethical approval of this study was obtained from the Research Ethics Committees of the Capital Region of Denmark. No participant suffered from neurological or hearing disorders. The pure-tone average of air conduction thresholds at 0.5, 1, 2 and 4 kHz (PTA4) were tested for hearing abilities and confirmed to be below 25 SPL HL.

B. Experimental Setup

The experimental setup consisted of five loudspeakers positioned around the participant at 1.2 m distance. The target loudspeaker was positioned 0° azimuth in front of the listener. The background noise, consisting of 4-talker babble, was presented from four loudspeakers located at $\pm 90^\circ$ and $\pm 150^\circ$ azimuth (16-talker babble noise in total). The spatial setup of the test is illustrated in Fig. 1.

Stimuli were routed through a sound card (RME Hammerfall DSB multiface II, Audio AG, Germany) and were played via loudspeakers KEF Q300 (KEF, United Kingdom).

A BioSemi ActiveTwo amplifier system (Biosemi, Netherlands) was used for EEG recording with a 64-channel cap mounted according to the international extended 10-20 system

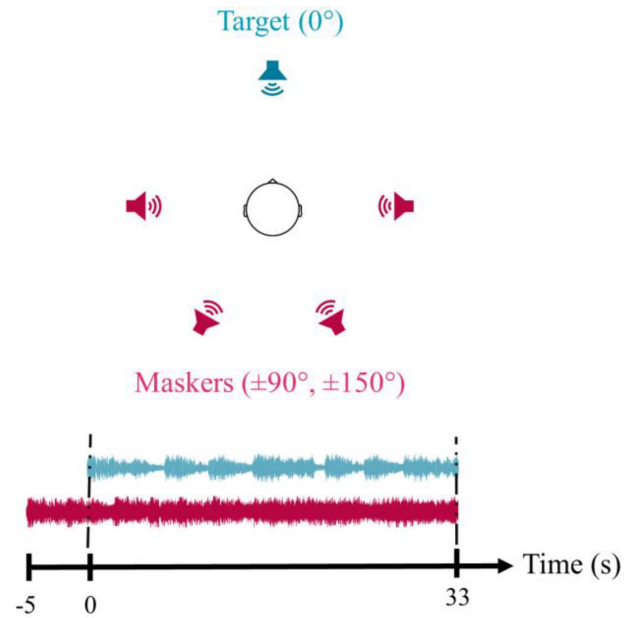


Fig. 1. The spatial setup of the experiment and the design of the test: Background noise was presented from the four loudspeakers (in red) for 38s. 5s after the onset of the background, the target was presented from the front loudspeaker (in blue) for 33s. At the end of each trial, the participants answered a two-choice question regarding the contents of the audio clip.

and a sampling frequency of 1024 Hz. The cap included DRL and

CMS electrodes as references for all other recording electrodes. Conductive gel was applied to the electrodes to obtain stable and below 50 mV offset voltage.

The ear-EEG recordings were acquired with a sampling rate of 1 KHz by a 32-channel portable TMSi MOBITA EEG amplifier (TMSi, Netherlands). In addition, the amplifier enabled active shielding (guarding) of the ear-electrodes all the way to the backside of each of the 12 electrodes. For each participant, earmould impressions were acquired in a session before the test in order to make them a personalized ear-EEG [33].

C. Design

Non-dramatic Danish news clips of neutral content were used for the target speech (33 s), as well as the background noise (38 s). Background noise consisted of 4-talker babble presented from each of the four loudspeakers in the back, resulting in a 16-talker babble noise.

The A-weighted sound pressure level produced by the babble was fixed at 70 dB overall (64 dB each) at the center of the loudspeakers where the participants were seated. The level of the target was varied across trials from 54–78 dB to generate four different SNRs: -16 , -8 , -4 and $+8$ dB. In this study, SNR was defined as the long-term average sound level of the target signal (with pauses longer than 200 ms being cut out) compared to the background noise.

There were 21 trials for each SNR, randomly distributed across 84 trials (with a rest period every 28 trials) and an

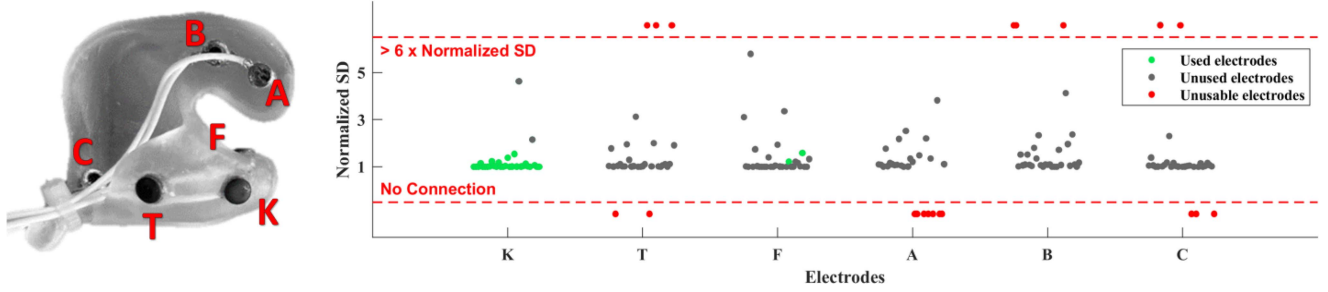


Fig. 2. Left: The location of the dry ear-EEG electrodes consisted of 6 electrodes in each ear (A, B, C, F, K, and T). Right: The normalized standard deviation (SD) for each electrode in ear-EEG for all participants' both ears (180 in total). The green electrodes (30) were used for further analysis (28 Ks and 2 Fs). The unusable electrodes below the "No Connection" or above "> 6 x Normalized SD" had no output or were too noisy, respectively.

additional 4 training trials in the beginning of the test. Each trial (Fig. 1) lasted 38 s while the 16-talker babble played in the background. The target speech was presented 5 s after the onset of the babble (i.e., after the baseline period) and then continued for 33 s, followed by a two-choice question about the content of the attended target audio clip. The percentage of correct answers were considered as performance accuracy.

D. Scalp EEG Pre-Processing

For the analysis of the scalp EEG, first power line noise was rejected with a 50-Hz notch filter with a quality factor of 25. Then a 3rd-order zero-phase Butterworth bandpass filter with cutoff frequencies of 1–40 Hz was applied to the data and the resulting signals were down-sampled to 256 Hz. Bad channels and trials were removed by visual inspection. On average 1.8 bad channels were detected per participant and interpolated using spline interpolation [34]. Also, 9.2% of the trials across all participants were rejected. No participant had more than 23.8% of trials rejected. The remaining trials were denoised using joint decorrelation method, as described in [35], [36].

For the resulting signals two different referencing methods were applied. The first method was the *common* average referencing which is acquired by subtracting the average of all channels from every single channel [22]. The second method was the *symmetrical* bi-polar referencing which is acquired by subtracting any two channels mirrored to each other in opposite hemispheres (e.g., T7 – T8). The latter referencing approach allowed us for a more direct comparison of the results of scalp and ear-EEG. The resulting signals were used for power extraction.

E. Ear-EEG Pre-Processing

For the analysis of the ear-EEG, power line noise was first rejected with a 50-Hz notch filter with a quality factor of 25. Then a 3rd-order zero-phase Butterworth bandpass filter with cutoff frequencies of 1–40 Hz was applied to the data. The resulting signals were downsampled to 256 Hz to have the same sampling frequency as scalp EEG, for any further comparison. In total, 5 trials were missed in the ear-EEG data due to trigger issues. In addition to those, bad trials were also removed by visual inspection which summed up to 10.8% of the trials across all

participants rejected. No participant had more than 27.2% of trials in ear-EEG data rejected.

The first step in using ear-EEG was to select the best pair of electrodes, one in each ear. Given that the quality of signals in each electrode varied from participant to participant, we decided to choose the most consistent electrodes based on their normalized standard deviation (SD). For this purpose, the SD of all the electrodes within each participant were normalized to the smallest SD in that participant (except when there was no connection for an electrode). Fig. 2 summarizes the normalized SD across all electrodes and participants (in both left and right ears). The further the normalized SD for each channel was from 1, the more chance that they were containing noise. Based on this, K electrode had the most consistent normalized SD and thus it was chosen as the main electrode for further analysis. However, based on the same metric, in two participants the quality for K electrode was poor and instead, F electrode was used in order to replace it. The F electrode is closely located to the K electrode, as both are placed close to the ear canal (Fig. 2 top panel).

After finding the best electrode pair for each participant (13 right-ear K-electrode paired with left-ear K-electrode, and 2 right-ear K-electrode paired with left-ear F-electrode), the right electrode was subtracted from the left one (similar to symmetrical referencing in the scalp EEG) and the resulting signal was used for power extraction in ear-EEG.

F. Alpha Power Extraction

To obtain time–frequency representation of EEG data, Morlet wavelets (7 cycles width) within the frequency range of 2 and 35 Hz, centered at 500 ms steps, were applied (using Fieldtrip toolbox [37]). Event-related spectral perturbation (ERSP) was then used to investigate how time-frequency data changed relative to baseline [38]. Positive ERSP is referred to as event-related synchronization (ERS) and negative ERSP is referred to as event-related desynchronization (ERD). The formula to calculate ERSP is as (1):

$$ERSP(t, f) = \frac{A(t, f) - \bar{R}(f)}{\bar{R}(f)} \times 100 \quad (1)$$

where A is the absolute power of the post-stimulus signal at time t and frequency f . \bar{R} is the absolute power of the baseline

signal averaged in time (-4 to 0 s), at the same frequency f . The first second was removed to avoid event-related potentials (ERP) due to the onset of the sound. Based on the visual inspection of averaged spectrograms across participants and SNRs in scalp EEG (common and symmetrical referencing) and ear EEG, we defined alpha power as 8–14 Hz averaged in parietal region (common referencing; CP1, CP2, CP3, CP4, CP5, CP6, CPz, P1, P2, P3, P4, P5, P6, P7, P8, Pz, PO3, PO4, PO7, PO8, POz) and temporal region (symmetrical referencing; FT7, FT8, T7, T8, TP7, TP8).

G. Linear Mixed Models

Linear mixed models (LMM) were implemented to analyze the effects of SNR on alpha modulation and performance. The applied models for statistical evaluation were as (2) and (3):

$$Alpha = \beta_0 + \beta_1 SNR + \beta_2 SNR^2 + \beta_3 SNR^3 + \mu_1 + \varepsilon \quad (2)$$

$$Perf = \beta_0 + \beta_1 SNR + \beta_2 SNR^2 + \beta_3 SNR^3 + \mu_1 + \varepsilon \quad (3)$$

In which β values are the weights of predictor (SNR) and μ_1 is the random effect of each participant, added with residual error (ε). Alpha power was extracted from scalp EEG (common and symmetrical referencing) and in ear-EEG. SNR values were centered around 0 (i.e., -11 , -3 , $+1$, $+13$ dB) in the model to avoid correlations between the linear, quadratic, and cubic effects [39]. The degree of the model was chosen based on the sigmoid pattern of grand average results of both alpha power and performance. The MATLAB syntax for (2) was as $Alpha \sim 1 + SNR^3 + (1|Subject)$ and for (3) as $Performance \sim 1 + SNR^3 + (1|Subject)$. The next model used was to predict alpha power based on performance with SNR as the random factor. This way, the model outcome will be generalizable to all SNRs and only be driven by performance, as in (4):

$$Alpha = \beta_0 + \beta_1 Performance + \mu_1 + \mu_2 + \varepsilon \quad (4)$$

In which μ_1 is the random effect of each participant and μ_2 is the random effect of SNR. The MATLAB syntax for (4) was as $Alpha \sim Performance + (1|Subject) + (1|SNR)$. In the Results section, the estimates (β) of the predictors, corresponding t-values and degrees of freedom (t_{DF}) and P-values for both models are reported.

H. Changes Over Time

One of the advantages of using continuous speech was the possibility of investigating changes of alpha power over time. For this purpose, the changes in slope of alpha power within 5-s time windows (no overlap) were investigated. To do so, first-degree polynomial curve was fit to the time-series data and LMM was used for evaluation of significant effect of first-order SNR on alpha slope in different time windows. This analysis was done on scalp EEG (parietal with common referencing and temporal with symmetrical referencing) and ear-EEG.

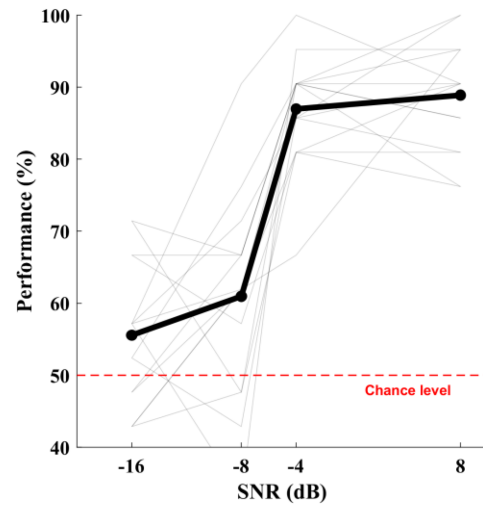


Fig. 3. Changes of performance with SNR, evaluated based on the questions regarding the contents of the target audio clip.

I. Correlation

In order to investigate if there is any correlation between performance (averaged for each condition) and alpha power (averaged for each condition), a robust Pearson skipped correlation was used to eliminate outliers by considering the data structure and mitigate the correlation bias of inter-dependency of samples within an individual [40]. The outcome of the skipped correlation was a bootstrapped data (1000 repetitions) with their 95% percentile confidence interval (CI). If the 95% CI did not contain zero, the Pearson coefficient r was considered significant. This analysis was done on scalp EEG (parietal with common referencing and temporal with symmetrical referencing) and ear-EEG.

The same method was used to look for the correlation between the temporal alpha in scalp EEG (symmetrical referencing) and the ear-EEG alpha power. This analysis was done to make sure that signals picked up by ear-EEG were closely correlated to the ones picked up by scalp EEG, around the ears.

III. RESULTS

A. Performance

Changes in performance showed that participants benefited from increasing SNR with a significant increase in accuracy ($\beta = 6.78$, $t_{56} = 0.96$, $p < 0.001$). There was no quadratic effect of SNR ($\beta = 0.01$, $t_{56} = 0.37$, $p = 0.710$), but rather a significant cubic effect ($\beta = -0.03$, $t_{56} = -5.42$, $p < 0.001$). The graph for performance accuracy is shown in Fig. 3.

B. Scalp EEG: Parietal - Common Reference

Applying LMM with common referencing for the scalp EEG in the parietal region showed a significant linear effect ($\beta = 2.68$, $t_{56} = 4.48$, $p = 0.007$), with no significant quadratic ($\beta = -0.04$, $t_{56} = -1.98$, $p = 0.052$), nor cubic effect ($\beta = -0.12$, $t_{56} = -1.91$, $p = 0.061$) of SNR on the alpha band. The

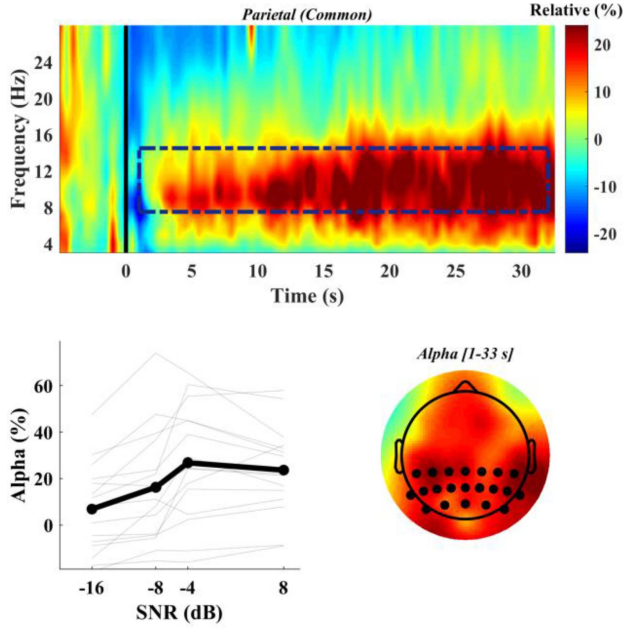


Fig. 4. Top row: The spectrogram of scalp EEG for the parietal electrodes, using common referencing. Bottom row left: Modulation of alpha power with SNR within the specified window. Bottom row right: Topographic map of alpha power during the stimuli.

graph for modulated alpha power, grand averaged spectrograms and corresponding topographic map are shown in Fig. 4.

C. Scalp EEG: Temporal - Symmetrical Reference

Applying LMM with symmetrical referencing for the scalp EEG in the temporal region revealed a significant linear effect ($\beta = 2.70$, $t_{56} = 2.85$, $p = 0.006$), and cubic effect ($\beta = -0.01$, $t_{56} = -2.16$, $p = 0.034$), but no significant quadratic effect of SNR ($\beta = -0.02$, $t_{56} = -1.27$, $p = 0.206$) on the alpha band. The graph for modulated alpha power, grand averaged spectrograms and corresponding topographic map are shown in Fig. 5.

D. Ear-EEG – Symmetrical Reference

In the ear-EEG, there was a significant linear effect ($\beta = 1.90$, $t_{56} = 3.00$, $p = 0.003$), and cubic effect ($\beta = -0.01$, $t_{56} = -2.31$, $p = 0.024$), but no significant quadratic effect of SNR ($\beta = -0.01$, $t_{56} = -1.22$, $p = 0.226$) on alpha band. Fig. 6 shows the modulated alpha power and grand average spectrograms of the ear-EEG data.

E. Alpha Changes Over Time

The evaluation of alpha slope within 5-s time windows revealed significant linear effect of SNR in the first 5 s of the stimuli in scalp EEG in the parietal (common referencing; $\beta = 0.41$, $t_{58} = 4.98$, $p < 0.001$) and temporal (symmetrical referencing; $\beta = 0.41$, $t_{58} = 4.56$, $p < 0.001$) regions, and also ear-EEG ($\beta = 0.29$, $t_{58} = 2.94$, $p = 0.004$). No other significant effect of SNR on slope was observed in any other time windows in any of the measures. Fig. 7 show the changes of alpha power over time and the slopes corresponding to each 5-s time windows.

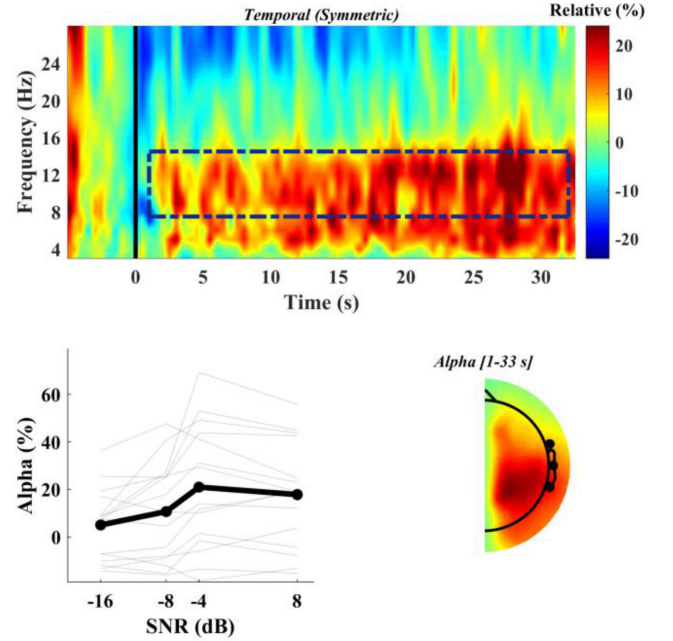


Fig. 5. Top row: The spectrogram of scalp EEG for the temporal electrodes, using symmetrical referencing. Bottom row left: Modulation of alpha power with SNR within the specified window. Bottom row right: Topographic map of alpha power during the stimuli.

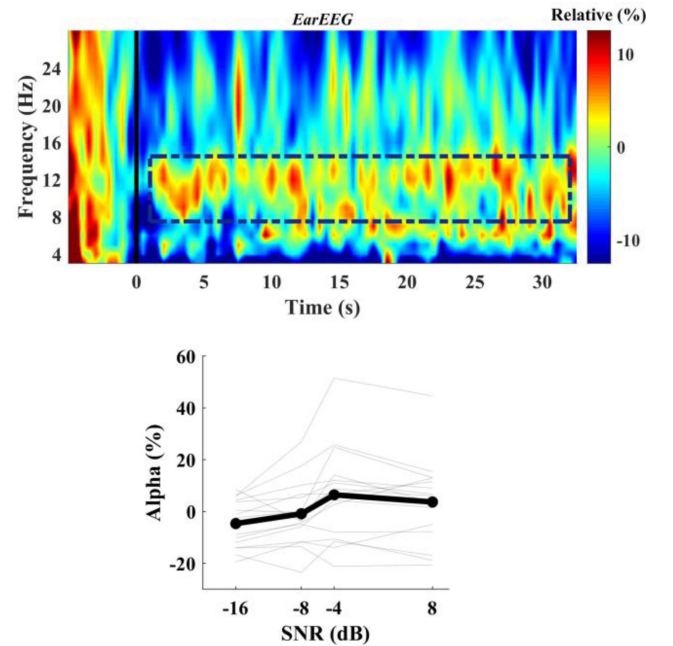


Fig. 6. Top row: The spectrogram of ear-EEG, using symmetrical referencing. Bottom row: Modulation of alpha power with SNR within the specified window.

F. Alpha Association With Performance

Predicting alpha power based on the performance in LMM, with SNR as random factor, showed significant effect of performance on parietal alpha (common referencing; $\beta = 0.27$, $t_{58} = 2.52$, $p = 0.014$), temporal alpha (symmetrical referencing;

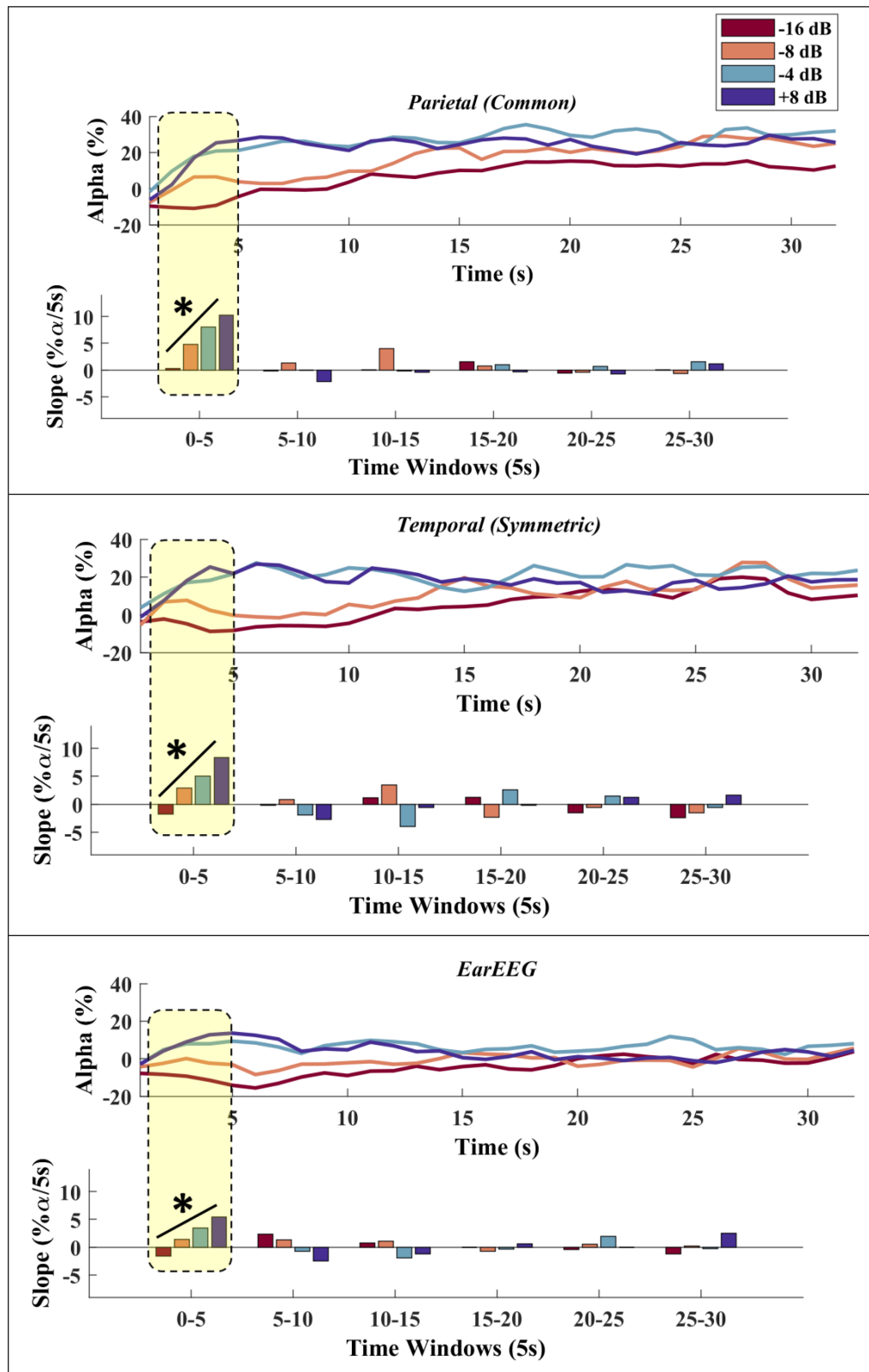


Fig. 7. Alpha slope in three different measurements. *Top panel*: scalp parietal using common referencing. *Middle panel*: temporal using symmetrical referencing. *Bottom panel*: ear-EEG. Upper plot within each panel shows alpha activity over time (moving average filter of 1s is applied for illustration purpose). Lower plot within each panel shows slope of alpha power in each 5-s window, using first degree polynomial fit. Significant linear trend corresponding to the time series are highlighted in a yellow box and an asterisk.

$\beta = 0.32$, $t_{58} = 4.31$, $p < 0.001$), and ear-EEG ($\beta = 0.21$, $t_{58} = 4.17$, $p < 0.001$).

In addition to that, Pearson skipped correlation also showed a significant correlation between alpha power and performance accuracy in scalp EEG in the parietal (common referencing; $r = 0.49$, 95% CI = [0.31, 0.67]) and temporal (symmetrical referencing; $r = 0.30$, 95% CI = [0.04, 0.53]) regions, and ear-EEG ($r = 0.30$, 95% CI = [0.04, 0.51]). The results are illustrated in Fig. 8.

G. Correlation Between Temporal EEG and Ear-EEG

Pearson skipped correlation showed significant correlation between alpha power extracted from temporal scalp EEG (symmetrical referencing) and ear-EEG ($r = 0.83$, 95% CI = [0.73, 0.89]). The results are shown in Fig. 9.

IV. DISCUSSION

A. Overview

Using four different SNRs, from very low to very high task demand, alpha power ERS decreased with increasing demand. This was present in both scalp EEG (using common or symmetrical referencing) and ear-EEG. Investigating the slope of alpha power over 33 s of the stimulus presentation in the same measures showed that there is a significant linear effect of SNR in the first 5 s.

Alpha ERS was also positively correlated to performance accuracy in scalp EEG and ear-EEG, which was based on answers to questions related to the content of the speech.

B. Alpha Modulation: A Marker for Effort or Performance?

Based on the literature, an increase in alpha power is often related to an increase in task demand of listening when short-sentence paradigms are used to evaluate effort [7], [8], [41], [42], [43]. In these paradigms, when the task is exceedingly difficult for individuals, an inverted U-shaped pattern of alpha power is observed, as participants disengage from the task [6], [41]. Studies with longer stimuli, however, often have reported the inverse effect (i.e., a decrease in alpha relating to an increase in demand) [11], [31], [32], [44], [45], [46].

Irrespective of referencing method (common or symmetrical), brain region (parietal or temporal), or EEG device (scalp or ear-EEG), the results of alpha power in this study showed a significant linear decrease with increasing demand (Figs. 4, 5, and 6). These results were in line with two of our previous studies, which showed less alpha ERS in more demanding continuous speech in hearing-impaired participants [31], [32]. One of the hypotheses on the functional role of alpha power is “gating by inhibition” [15], [17]. This theory postulates that alpha ERS shuts down irrelevant task regions which helps routing the information to the task-relevant areas. Based on this theory, if alpha power is a measure of listening effort, it should increase with increasing task demand (i.e., lower SNR). Additionally, since the task demand varied from very easy to very difficult in this study (24 dB SNR span), an inverted U-shaped

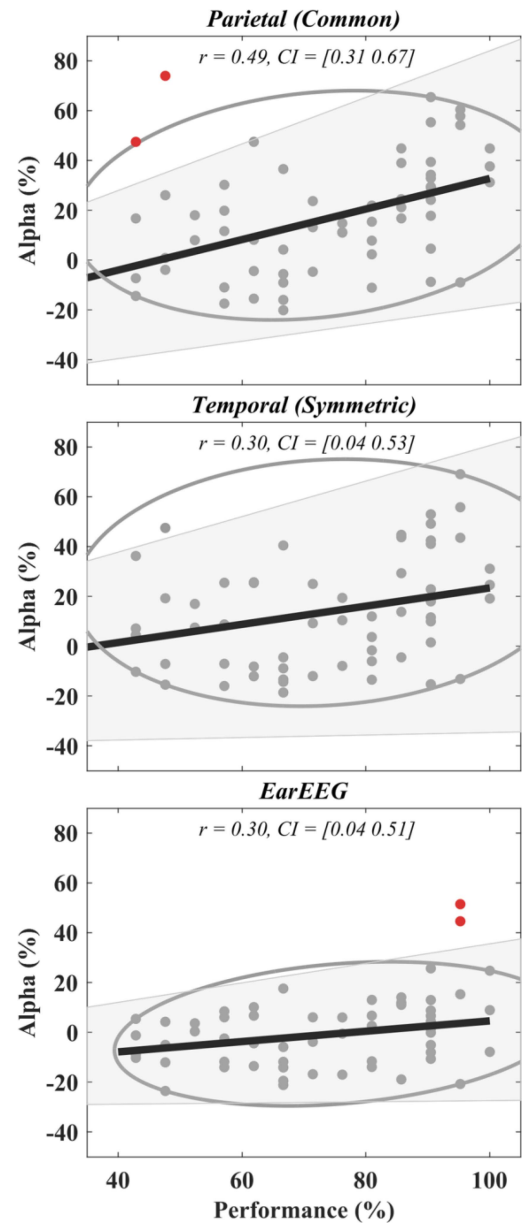


Fig. 8. Skipped Pearson correlation (r) between performance accuracy and alpha power in scalp parietal using common referencing (top plot), temporal using symmetrical referencing (middle plot) and ear-EEG (bottom plot). The gray ellipses are the estimated area by the skipped correlation based on the structure of the data to determine if a data sample is an outlier or not. Data samples outside the gray ellipses (red dots) were considered as outliers and left out for r estimation. The gray area shows 95% CIs of the bootstrapped data which were positively significant for all three measurements.

pattern of alpha power over SNR was expected. However, the results showed no inverted U-shaped alpha power. Instead, alpha power increased in sigmoidal fashion with increasing SNR (i.e., opposite to “gating by inhibition” theory). This may indicate that alpha power shows something other than listening effort.

In a recent study [46] on vocoded continuous speech, increased demand (i.e., greater speech degradation) led to decreased alpha power as well. The authors suggested that during a

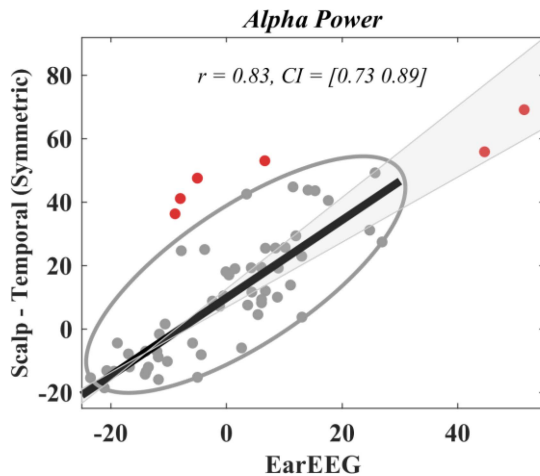


Fig. 9. Skipped Pearson correlation (r) between alpha power extracted from temporal scalp EEG (symmetrical referencing) and ear-EEG. The gray ellipses are the estimated area by the skipped correlation based on the structure of the data to determine if a data sample is an outlier or not. Data samples outside the gray ellipses (red dots) were considered as outliers and left out for r estimation. The gray area shows 95% CIs of the bootstrapped data which were positively significant.

continuous speech task, which might require complex linguistic processing in the brain, increased ability to track and understand speech is accompanied by increased alpha [46]. In fact, our results also showed that alpha during continuous speech might be a better indicator of speech intelligibility than listening effort. Better performance led to increased alpha ERS during the task. Performance accuracy could also predict alpha power, without the influence of SNR on performance, in the statistical model. While performance accuracy in this study was a simplistic measure of speech intelligibility, it still significantly benefitted from increasing SNR. This provides further evidence against the idea that alpha reflects listening effort during continuous speech. While listening effort has been shown to change in an inverted U-shaped pattern (e.g., [5], [6], [47]), speech recognition is best described as sigmoidal with changes of task demand [48], [49]. Therefore, it is unlikely that one measure can reflect both at the same time, at least when such a wide range of task demand (from -16 dB to $+8$ dB SNR) was used.

However, the problem in interpreting the correlation between performance and alpha ERS is that the causality direction between them is unknown. It is unclear whether increased alpha activity led to better speech intelligibility (“top-down” attention), or increased SNR led to better speech intelligibility as well as entrainment of alpha power in the brain (“bottom-up” attention). Therefore, based on these results, the role of alpha power in speech processing is unclear. We can only conclude that alpha power is correlated to performance which is an indirect measure of speech intelligibility.

C. Changes of Alpha Over Continuous Speech

One of the advantages of using long, continuous speech is to explore changes of alpha activity over an extended period of listening. Reasonably, it cannot be expected that a listener invests

effort continuously at a constant level during the whole presentation of continuous speech [50]. A person can adapt to specific listening difficulties, get fatigued, or lose/gain motivation over time.

Using continuous speech that lasted for 33 s, the slope of alpha power in the first 5 s showed a linear trend with SNR in both scalp (common or symmetric referencing) and ear-EEG. In other words, higher SNRs had larger leap from alpha ERS to ERS in the first 5 s of the stimuli. The early transition from alpha ERS to ERS might be an early indicator of whether a listening situation is easily intelligible or not.

The importance of this finding is that in such a continuous-speech paradigm, only the first 5 s of the stimulus presentation can be still informative about the intelligibility manipulations. However, it is important to note that while the averaged alpha power over 33 s showed cubic effect as well, alpha slope within the first 5 s only showed linear effect of SNR. Therefore, using averaged alpha over the whole stimuli may correspond better to the performance, which also followed a cubic pattern with SNR.

D. Feasibility of Ear-EEG to Measure Alpha Power

Ear-EEG is a feasible method for ambulatory measurement of EEG. While ear-EEG is limited to picking up the electrical brain signals around the ears, it has been shown that it can be successfully used in auditory studies to decode the attended talker in the presence of distracting talkers [28], [29], [30]. However, to our knowledge no studies have investigated alpha oscillations in continuous speech using ear-EEG.

The first issue of using ear-EEG in a realistic listening scenario (i.e., continuous speech) is the quality of recorded data compared to the scalp EEG. In this study, for scalp EEG, conductive gel was applied for better contact to the skin, but for ear-EEG the electrodes were used dry. While using conductive gel leads to better-quality signals in general, it is important to implement dry electrodes for any further application of ear-EEGs in real life. Even with the differences in impedance of electrodes, ear-EEG signals had good quality compared to scalp-EEG signals (on average 1.3 more trials were rejected in ear-EEG compared to scalp EEG), especially using the selected electrodes of the ear-EEG (e.g., K and F, refer to Fig. 2). The high correlation in alpha power between the ear-EEG (using dry electrodes) and the temporal region of scalp EEG (using wet electrodes) was another indication that the quality of recorded data in ear-EEG was close to signals from scalp EEG (refer to Fig. 9).

The second aspect of the ear-EEG results was how the changes in alpha power with task demand could be compared in scalp EEG and ear-EEG. We could see that alpha power changed in a sigmoidal pattern in both scalp EEG (temporal, symmetrical referencing) and ear-EEG by manipulating SNR in continuous speech and were correlated to the performance. Similar to scalp EEG, we speculate that alpha power during elicited by such stimuli is less a reflection of listening effort and more a measure of performance. The slope of alpha power within the first 5 s of the stimuli also showed similar pattern in the ear-EEG compared to scalp EEG (all showed linear effect of SNR on the alpha slope in the first 5 s).

More research is required to investigate the feasibility of measuring performance or speech intelligibility using ear-EEG, a question that has potential for real-world impact if implemented in hearing aids. To this aim, using real-life scenarios similar to the conversation-like stimuli used in this study is encouraged.

V. CONCLUSION

Wearable technology such as ear-EEG can give us insight into brain activity in everyday environments. To our knowledge, this is the first study that investigates alpha oscillations in ear-EEG during effortful continuous speech. To do this, we recorded brain signals with a 64-channel scalp EEG simultaneously with dry ear-EEG electrodes during a continuous speech-in-noise task. Four different SNRs were generated to manipulate task demand. The results showed decreasing SNR (i.e., increasing task demand) led to decreasing alpha power in both scalp EEG and ear-EEG. Also, the slope of alpha power within the first 5 s of the speech showed similar linear pattern in scalp EEG and ear-EEG with increasing task demand. We report that ear-EEG can measure changes in performance using alpha activity as a marker, providing an indirect measure of speech intelligibility.

ACKNOWLEDGMENT

The authors would like to thank Allan Theill Sørensen, Eline Borch Petersen, and Hjørdis Huginsdóttir Eide for their contribution to the study by preparing the experiment and collecting the data.

REFERENCES

- [1] M. K. Pichora-Fuller et al., "Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL)," *Ear Hear.*, vol. 37, no. 1, pp. 5S–27S, 2016, doi: [10.1097/AUD.0000000000000312](https://doi.org/10.1097/AUD.0000000000000312).
- [2] J. E. Peelle, "Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior," *Ear Hear.*, vol. 39, no. 2, pp. 204–214, 2018, doi: [10.1097/AUD.0000000000000494](https://doi.org/10.1097/AUD.0000000000000494).
- [3] D. Wendt, R. K. Hietkamp, and T. Lunner, "Impact of noise and noise reduction on processing effort," *Ear Hear.*, vol. 38, no. 6, pp. 690–700, 2017, doi: [10.1097/AUD.0000000000000454](https://doi.org/10.1097/AUD.0000000000000454).
- [4] T. Koelewijn et al., "Pupil dilation uncovers extra listening effort in the presence of a single-talker masker," *Ear Hear.*, vol. 33, no. 2, pp. 291–300, 2012, doi: [10.1097/AUD.0b013e3182310019](https://doi.org/10.1097/AUD.0b013e3182310019).
- [5] B. Ohlenforst et al., "Impact of SNR, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response," *Hear. Res.*, vol. 365, pp. 90–99, 2018, doi: [10.1016/j.heares.2018.05.003](https://doi.org/10.1016/j.heares.2018.05.003).
- [6] L. Decruy et al., "Top-down modulation of neural envelope tracking: The interplay with behavioral, self-report and neural measures of listening effort," *Eur. J. Neurosci.*, vol. 52, no. 5, pp. 3375–3393, 2020, doi: [10.1111/ejn.14753](https://doi.org/10.1111/ejn.14753).
- [7] M. G. Wisniewski, E. R. Thompson, and N. Iyer, "Theta- and alpha-power enhancements in the electroencephalogram as an auditory delayed match-to-sample task becomes impossibly difficult," *Psychophysiology*, vol. 54, no. 12, pp. 1916–1928, 2017, doi: [10.1111/psyp.12968](https://doi.org/10.1111/psyp.12968).
- [8] J. Obleser et al., "Adverse listening conditions and memory load drive a common alpha oscillatory network," *J. Neurosci.*, vol. 32, no. 36, pp. 12376–12383, 2012, doi: [10.1523/jneurosci.4908-11.2012](https://doi.org/10.1523/jneurosci.4908-11.2012).
- [9] M. G. Wisniewski et al., "Sustained frontal midline theta enhancements during effortful listening track working memory demands," *Hear. Res.*, vol. 358, pp. 37–41, 2017, doi: [10.1016/j.heares.2017.11.009](https://doi.org/10.1016/j.heares.2017.11.009).
- [10] A. Dimitrijevic et al., "Neural indices of listening effort in noisy environments," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, 2019, doi: [10.1038/s41598-019-47643-1](https://doi.org/10.1038/s41598-019-47643-1).
- [11] C. M. McMahon et al., "Monitoring alpha oscillations and pupil dilation across a performance-intensity function," *Front. Psychol.*, vol. 7, 2016, Art. no. 745, doi: [10.3389/fpsyg.2016.00745](https://doi.org/10.3389/fpsyg.2016.00745).
- [12] D. Pinal, M. Zurrón, and F. Díaz, "Effects of load and maintenance duration on the time course of information encoding and retrieval in working memory: From perceptual analysis to post-categorization processes," *Front. Hum. Neurosci.*, vol. 8, 2014, Art. no. 165, doi: [10.3389/fnhum.2014.00165](https://doi.org/10.3389/fnhum.2014.00165).
- [13] F. Kretschmar et al., "Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media," *PLoS One*, vol. 8, no. 2, Art. no. e56178, 2013, doi: [10.1371/journal.pone.0056178](https://doi.org/10.1371/journal.pone.0056178).
- [14] O. Jensen et al., "Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task," *Cereb. Cortex*, vol. 12, no. 8, pp. 877–882, 2002.
- [15] W. Klimesch, "Alpha-band oscillations, attention, and controlled access to stored information," *Trends Cogn. Sci.*, vol. 16, no. 12, pp. 606–617, 2012, doi: [10.1016/j.tics.2012.10.007](https://doi.org/10.1016/j.tics.2012.10.007).
- [16] N. Weisz and J. Obleser, "Synchronisation signatures in the listening brain: A perspective from non-invasive neuroelectrophysiology," *Hear. Res.*, vol. 307, pp. 16–28, 2014, doi: [10.1016/j.heares.2013.07.009](https://doi.org/10.1016/j.heares.2013.07.009).
- [17] O. Jensen and A. Mazaheri, "Shaping functional architecture by oscillatory alpha activity: Gating by inhibition," *Front. Hum. Neurosci.*, vol. 4, 2010, Art. no. 186, doi: [10.3389/fnhum.2010.00186](https://doi.org/10.3389/fnhum.2010.00186).
- [18] A. Dimitrijevic et al., "Cortical alpha oscillations predict speech intelligibility," *Front. Hum. Neurosci.*, vol. 11, 2017, Art. no. 88, doi: [10.3389/fnhum.2017.00088](https://doi.org/10.3389/fnhum.2017.00088).
- [19] A. Sarampalis et al., "Objective measures of listening effort: Effects of background noise and noise reduction," *J. Speech, Lang. Hear. Res.*, vol. 52, no. 5, pp. 1230–1240, 2009, doi: [10.1044/1092-4388\(2009\)08-0111](https://doi.org/10.1044/1092-4388(2009)08-0111).
- [20] R. Houben, M. Van Doorn-Bierman, and W. A. Dreschler, "Using response time to speech as a measure for listening effort," *Int. J. Audiology*, vol. 52, no. 11, pp. 753–761, 2013, doi: [10.3109/14992027.2013.832415](https://doi.org/10.3109/14992027.2013.832415).
- [21] T. Lunner et al., "Using speech recall in hearing aid fitting and outcome evaluation under ecological test conditions," *Ear Hear.*, vol. 37, no. 1, pp. 145S–154S, 2016, doi: [10.1097/AUD.0000000000000294](https://doi.org/10.1097/AUD.0000000000000294).
- [22] S. J. Luck, *An Introduction to the Event-Related Potential Technique*, 2nd ed. Cambridge, MA, USA: MIT Press, 2014.
- [23] D. Looney et al., "The in-the-ear recording concept: User-centered and wearable brain monitoring," *IEEE Pulse*, vol. 3, no. 6, pp. 32–42, Nov. 2012, doi: [10.1109/MPUL.2012.2216717](https://doi.org/10.1109/MPUL.2012.2216717).
- [24] M. G. Bleichner et al., "Exploring miniaturized EEG electrodes for brain-computer interfaces. An EEG you do not see?," *Physiol. Rep.*, vol. 3, no. 4, 2015, Art. no. e12362, doi: [10.14814/phys2.12362](https://doi.org/10.14814/phys2.12362).
- [25] S. L. Kappel et al., "Dry-contact electrode ear-EEG," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 150–158, Jan. 2019, doi: [10.1109/TBME.2018.2835778](https://doi.org/10.1109/TBME.2018.2835778).
- [26] K. B. Mikkelsen et al., "EEG recorded from the ear: Characterizing the ear-EEG method," *Front. Neurosci.*, vol. 9, 2015, Art. no. 438, doi: [10.3389/fnins.2015.00438](https://doi.org/10.3389/fnins.2015.00438).
- [27] J. W. Ahn et al., "Wearable in-the-ear EEG system for SSVEP-based brain-computer interface," *Electron. Lett.*, vol. 54, no. 7, pp. 413–414, 2018, doi: [10.1049/el.2017.3970](https://doi.org/10.1049/el.2017.3970).
- [28] L. Fiedler et al., "Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech," *J. Neural Eng.*, vol. 14, no. 3, 2017, Art. no. 036020, doi: [10.1088/1741-2552/aa66dd](https://doi.org/10.1088/1741-2552/aa66dd).
- [29] B. Mirkovic et al., "Target speaker detection with concealed EEG around the ear," *Front. Neurosci.*, vol. 10, 2016, Art. no. 349, doi: [10.3389/FNINS.2016.00349](https://doi.org/10.3389/FNINS.2016.00349).
- [30] M. G. Bleichner, B. Mirkovic, and S. Debener, "Identifying auditory attention with ear-EEG: CEEGrid versus high-density cap-EEG comparison," *J. Neural Eng.*, vol. 13, no. 6, Oct. 2016, Art. no. 066004, doi: [10.1088/1741-2560/13/6/066004](https://doi.org/10.1088/1741-2560/13/6/066004).
- [31] T. S. Ala et al., "An exploratory study of EEG alpha oscillation and pupil dilation in hearing-aid users during effortful listening to continuous speech," *PLoS One*, vol. 15, no. 7, Jul. 2020, Art. no. e0235782, [Online]. Available: <https://doi.org/10.1371/journal.pone.0235782>
- [32] L. Fiedler et al., "Hearing aid noise reduction lowers the sustained listening effort during continuous speech in noise—A combined pupillometry and EEG study," *Ear Hear.*, vol. 42, no. 6, pp. 1590–1601, 2021, doi: [10.1097/AUD.0000000000001050](https://doi.org/10.1097/AUD.0000000000001050).
- [33] K. B. Mikkelsen et al., "Accurate whole-night sleep monitoring with dry-contact ear-EEG," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 16824, doi: [10.1038/s41598-019-53115-3](https://doi.org/10.1038/s41598-019-53115-3).

- [34] F. Perrin et al., "Mapping of scalp potentials by surface spline interpolation," *Electroencephalography Clin. Neurophysiol.*, vol. 66, no. 1, pp. 75–81, 1987, doi: [10.1016/0013-4694\(87\)90141-6](https://doi.org/10.1016/0013-4694(87)90141-6).
- [35] A. De Cheveigné and L. C. Parra, "Joint decorrelation, a versatile tool for multichannel data analysis," *NeuroImage*, vol. 98, pp. 487–505, 2014, doi: [10.1016/j.neuroimage.2014.05.068](https://doi.org/10.1016/j.neuroimage.2014.05.068).
- [36] A. de Cheveigné and D. Arzounian, "Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data," *Neuroimage*, vol. 172, pp. 903–912, 2018, doi: [10.1016/j.neuroimage.2018.01.035](https://doi.org/10.1016/j.neuroimage.2018.01.035).
- [37] R. Oostenveld et al., "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, 2011, Art. no. 156869, doi: [10.1155/2011/156869](https://doi.org/10.1155/2011/156869).
- [38] G. Pfurtscheller, "Functional brain imaging based on ERD/ERS," *Vis. Res.*, vol. 41, no. 10/11, pp. 1257–1260, 2001, doi: [10.1016/S0042-6989\(00\)00235-2](https://doi.org/10.1016/S0042-6989(00)00235-2).
- [39] X. A. Harrison et al., "A brief introduction to mixed effects modelling and multi-model inference in ecology," *PeerJ*, vol. 6, 2018, Art. no. e4794, doi: [10.7717/peerj.4794](https://doi.org/10.7717/peerj.4794).
- [40] C. R. Pernet, R. Wilcox, and G. A. Rousselet, "Robust correlation analyses: False positive and power validation using a new open source matlab toolbox," *Front. Psychol.*, vol. 3, 2013, Art. no. 606, doi: [10.3389/fpsyg.2012.00606](https://doi.org/10.3389/fpsyg.2012.00606).
- [41] E. B. Petersen et al., "Hearing loss impacts neural alpha oscillations under adverse listening conditions," *Front. Psychol.*, vol. 6, 2015, Art. no. 177, doi: [10.3389/fpsyg.2015.00177](https://doi.org/10.3389/fpsyg.2015.00177).
- [42] M. Wöstmann, S. J. Lim, and J. Obleser, "The human neural alpha response to speech is a proxy of attentional control," *Cereb. Cortex*, vol. 27, no. 6, pp. 3307–3317, 2017, doi: [10.1093/cercor/bhx074](https://doi.org/10.1093/cercor/bhx074).
- [43] M. Wöstmann et al., "Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits," *J. Neurosci.*, vol. 35, no. 4, pp. 1458–1467, 2015, doi: [10.1523/JNEUROSCI.3250-14.2015](https://doi.org/10.1523/JNEUROSCI.3250-14.2015).
- [44] K. Miles et al., "Objective assessment of listening effort: Coregistration of pupillometry and EEG," *Trends Hear.*, vol. 21, 2017, Art. no. 2331216517706396, doi: [10.1177/2331216517706396](https://doi.org/10.1177/2331216517706396).
- [45] J. Hjortkjær et al., "Cortical oscillations and entrainment in speech processing during working memory load," *Eur. J. Neurosci.*, vol. 51, no. 5, pp. 1279–1289, 2020, doi: [10.1111/ejn.13855](https://doi.org/10.1111/ejn.13855).
- [46] A. Hauswald et al., "Degradation levels of continuous speech affect neural speech tracking and alpha power differently," *Eur. J. Neurosci.*, vol. 55, no. 11/12, pp. 3288–3302, Aug. 2020, doi: [10.1111/ejn.14912](https://doi.org/10.1111/ejn.14912).
- [47] D. Wendt et al., "Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test," *Hear. Res.*, vol. 369, pp. 67–78, 2018, doi: [10.1016/j.heares.2018.05.006](https://doi.org/10.1016/j.heares.2018.05.006).
- [48] J. Festen and R. Plomp, "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Amer.*, vol. 88, no. 4, pp. 1725–1736, 1990, doi: [10.1121/1.400247](https://doi.org/10.1121/1.400247).
- [49] J. B. Nielsen and T. Dau, "Development of a Danish speech intelligibility test," *Int. J. Audiology*, vol. 48, no. 10, pp. 729–741, 2009, doi: [10.1080/14992020903019312](https://doi.org/10.1080/14992020903019312).
- [50] D. J. Strauss and A. L. Francis, "Toward a taxonomic model of attention in effortful listening," *Cogn. Affect. Behav. Neurosci.*, vol. 17, no. 4, pp. 809–825, Aug. 2017, doi: [10.3758/s13415-017-0513-0](https://doi.org/10.3758/s13415-017-0513-0).