Introduction

In this project, we were tasked with analysing EEG data recorded during a study investigating the function of rhythm in language processing. Specifically, the study investigated the role of rhythm in helping to process sentence ambiguity and if the age at which the participant acquired the language was influential.

A group of 45 participants in 3 groups (monolinguals of German, Turkish early learners of German & Turkish late learners of German) were presented with rhythmically regular/irregular subject/object first sentences in a within-subject comparison setup. The experimental steps were as shown in figure 1. Each time a participant heard a sentence, he/she was presented with a comprehension task and asked to determine if a written rephrase of what they saw matched what they heard.

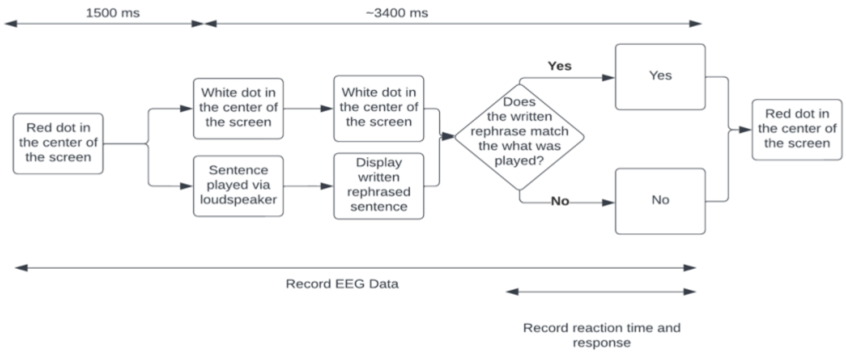

Figure 1: Experimental procedure

EEG recordings were collected from 59 electrode sites at a sampling frequency of 500 Hz. The study focused on 4 scalp distributions of interest as shown in table 1. All channels were processed in our analysis.

Table 1: Scalp regions of interest and corresponding channels.

Region	Channel
Left anterior	F1, F3, F5, FC1, FC3, FC5
Right anterior	F2, F4, F6, FC2, FC4, FC6
Left posterior	CP1, CP3, CP5, P1, P3, P5
Right posterior	CP2, CP4, CP6, P2, P4, P6

We conducted a time-resolved frequency analysis on the data collected from the late learners of the German group (n = 15) presented with regular rhythm sentences to investigate the contrast in activity between non-ambiguous (NA) (i.e., control) and ambiguous (AM) (i.e., experimental) conditions in the theta band (4-8 Hz). The experimental condition and control condition data were 15 x 44 x 64 x 751 and 15 x 44 x 64 x 276 (participants x trials x channels x samples) respectively. The EEG reading from each trial was baselined by subtracting the mean of the baseline time (i.e., the points till time t = 0) from all data points. An example of an ERP of a random participant (participant 12), averaged across all trials, for a random channel (channel 30) is shown in figure 2.

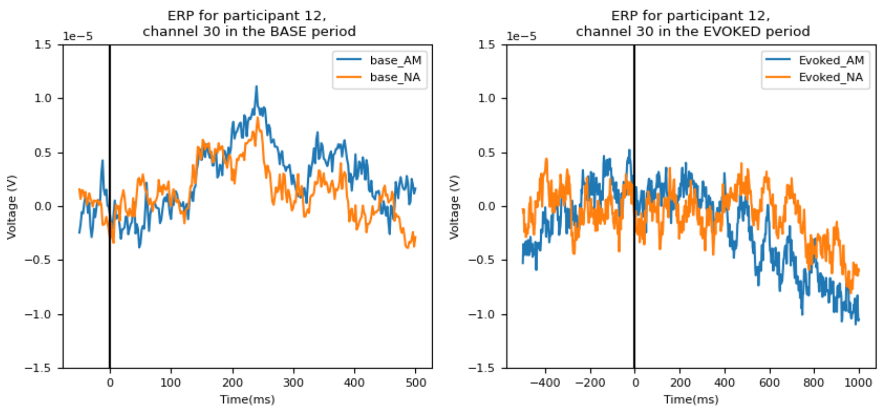

Figure 2: Baselined ERP example for a random participant.

Figure 3 shows the average ERP across all participants for a random channel (channel 30).

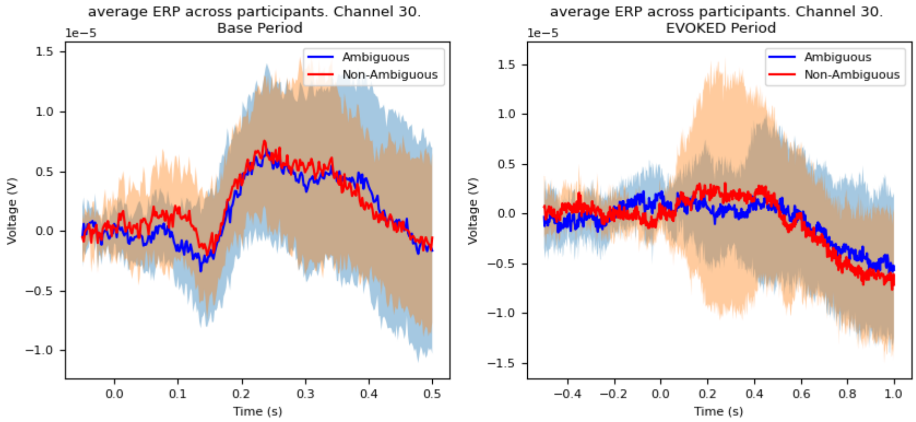

Figure 3: Average ERP for the entire group on a random channel. Shaded region indicates confidence interval

Time-Resolved Frequency Analysis

The necessary constants used for performing the frequency analysis are shown in table 3. Based on these values, the number of frequencies that can be analyzed at the given resolution was 126.

Table 2 : Necessary constants performing FFT in the Theta band.

Constant	Description	Value
srate	Sampling rate	500 Hz
nyquist_freq	Nyquist frequency : The maximum analyzable frequency at this sampling rate	250 Hz
freq_res	The frequency resolution for the Theta band (4-8Hz)	2 Hz
N_samples_FFT	The number of samples required to perform FFT	250

First, we extracted the necessary samples to create a baseline spectrum (for both control and experimental conditions) using 1 window of length 250 for each participant, trial, and channel. Next, we carried out a windowing procedure (on both experimental & control data) to convert each signal of length 751 to a windowed signal. Using the constants in table 2, this resulted in each signal being transformed from a one-dimensional array of 1x751 samples to a windowed signal of 21 windows where each window was 250 samples long. Finally, to each of these signal windows, we applied a FFT to extract the power of each frequency for a given window. The identical process of applying the FFT was applied to the baseline samples previously extracted (except in this case it was only 1 window of length 250). In each case, the result is a power spectrum for each window of 126 frequencies. An example of the results of this on one window is shown in figure 4.

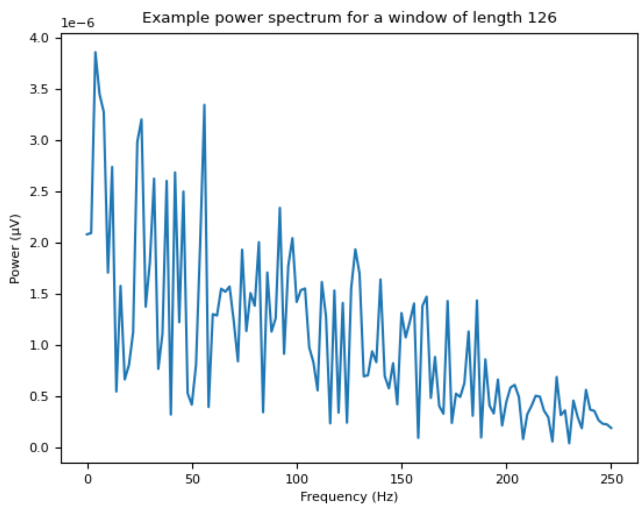

Figure 4: After applying the windowing procedure to each signal, each signal is analyzed for the power of a given frequency.

Figure 5 shows several plots demonstrating the output from the above procedure on the baseline (for both experimental and control). Each plot has been truncated on the x-axis at a length of 100 to focus more the desired area as neural signals show a 1/f exponential decay relationship between power and frequency.

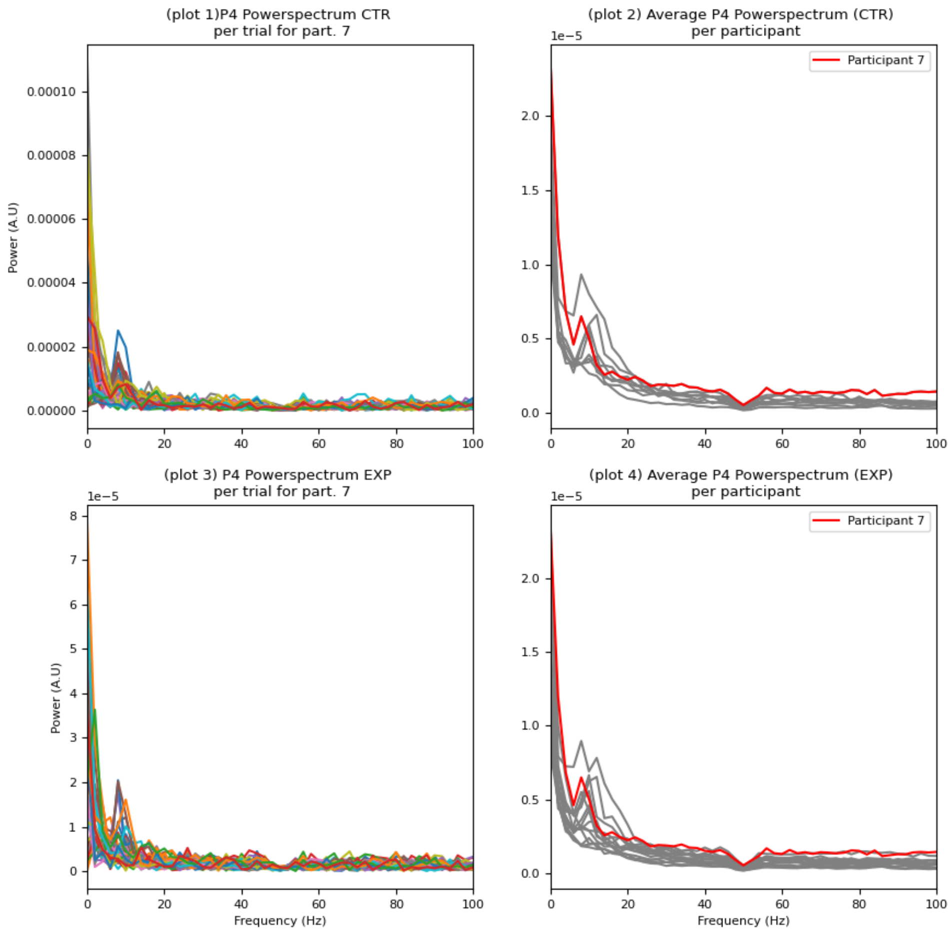

Figure 5: Plots on the left column show all baseline power spectrums for all trials, on one channel for one participant. Plots on the right column show average over all trials for one participant (red) and all other participants (grey) on one channel (i.e., calculating the mean power at each frequency across multiple trails)

Next, we performed a baseline normalization of all the power spectrums in the evoked period by dividing each window by the baseline power spectrum for each participant per trial per channel. Figure 6 shows an example of the normalized TFR for a single channel for a random participant.

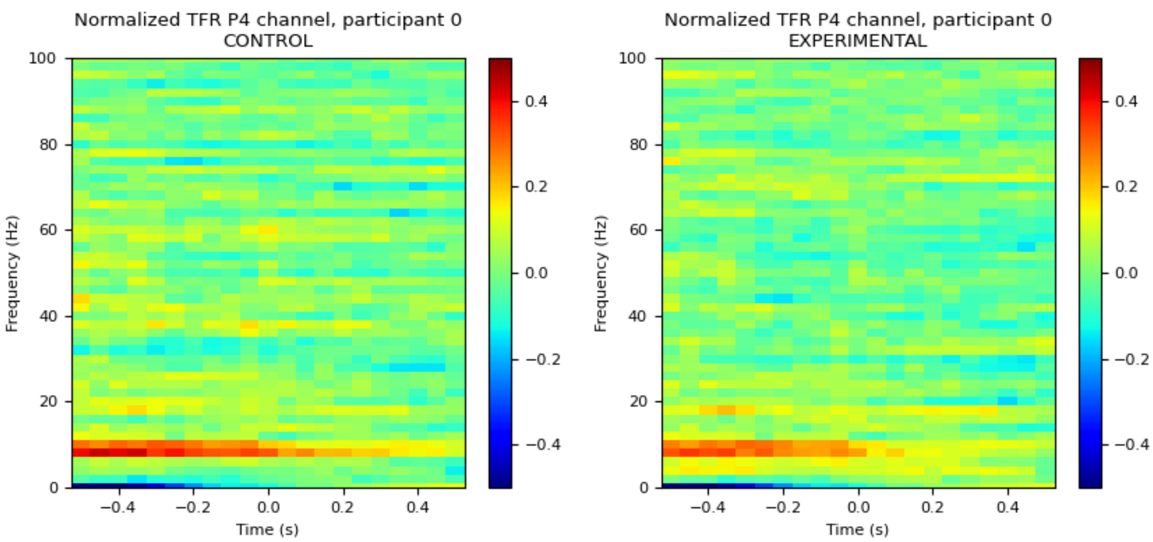

Figure 6: An example TFR after each window of analysis has been normalized by the baseline power at each frequency

Analysis & Discussion

To investigate the difference between the experimental and control conditions, we constructed a contrast of the normalized evoked TFRs by subtracting the experimental TFRs from the control TFRs. We tested for significant differences by conducting a one-sample t-test in the 200 ms after the onset of the critical item. An example for one channel (Fpz) is shown in figure 7 and 8. As shown by the p-value in the figure, for this example, there is no statistically significant difference between experimental and control conditions.

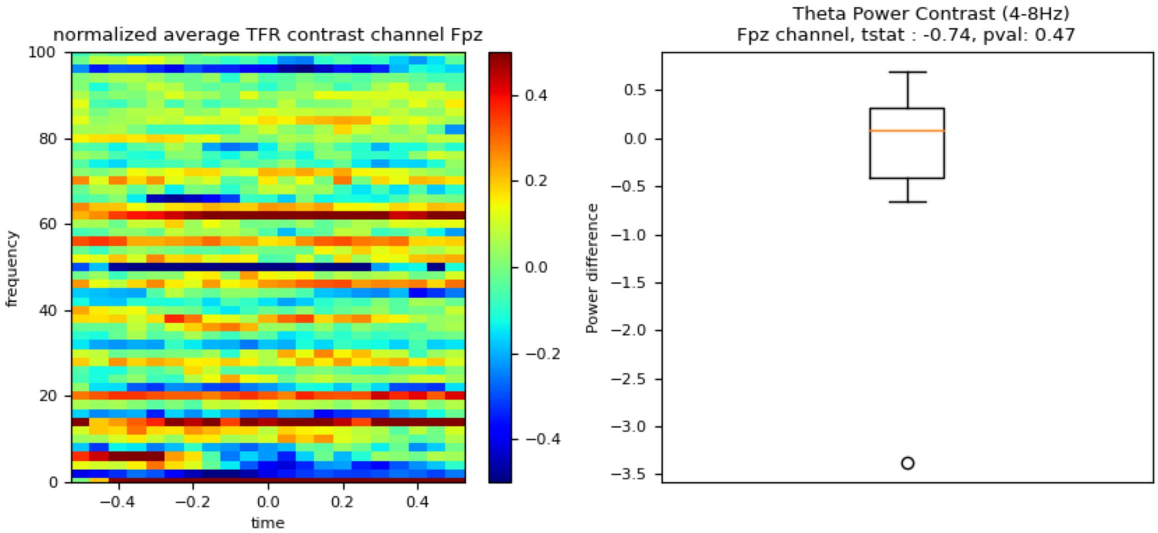

Figure 7: An example TFR contrast for one channel. Testing for the hypothesis that there is a difference between the control and experimental condition for this channel.

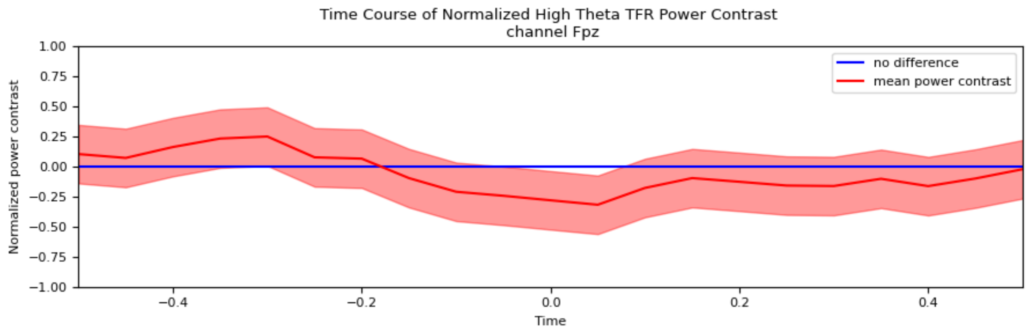

Figure 8: Shows the variation of mean-power contrast around 0 for channel Fpz.

We investigated the regions of interest for the study (Left anterior (fig 9), right anterior (fig 10), left posterior(fig 11) and right posterior (fig 12)), defined in table 1 for any significant difference in activity in the Theta band.

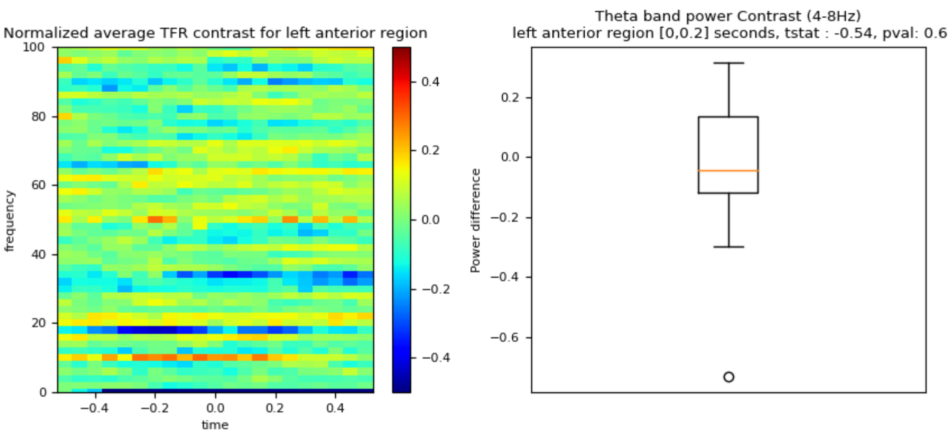

Figure 9: Testing for any statistically significant difference between experimental and control condition in contrast for the left anterior region in the time interval of [0,0.2] seconds.

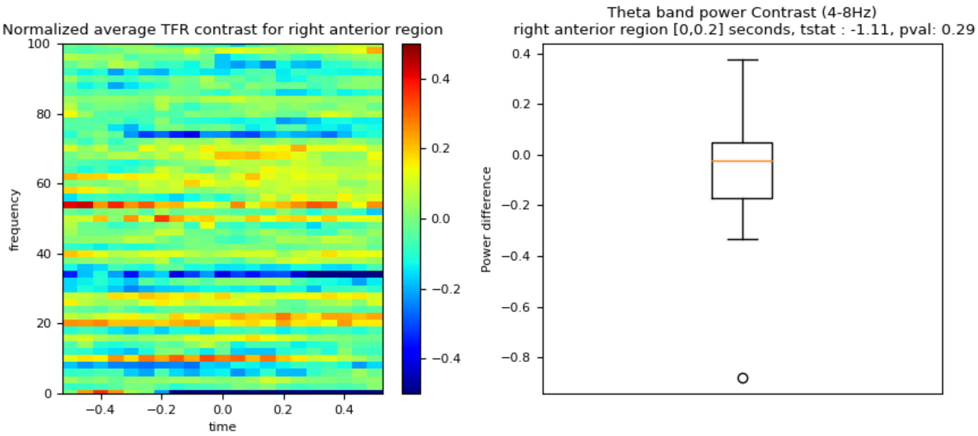

Figure 10: Testing for any statistically significant difference between experimental and control condition in contrast for the right anterior region in the time interval of [0,0.2] seconds.

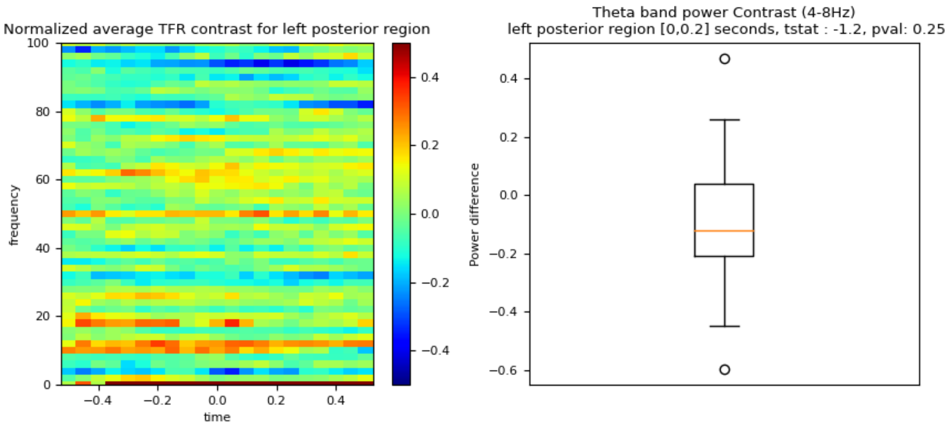

Figure 11: Testing for any statistically significant difference between experimental and control condition in contrast for the left posterior region in the time interval of [0,0.2] seconds.

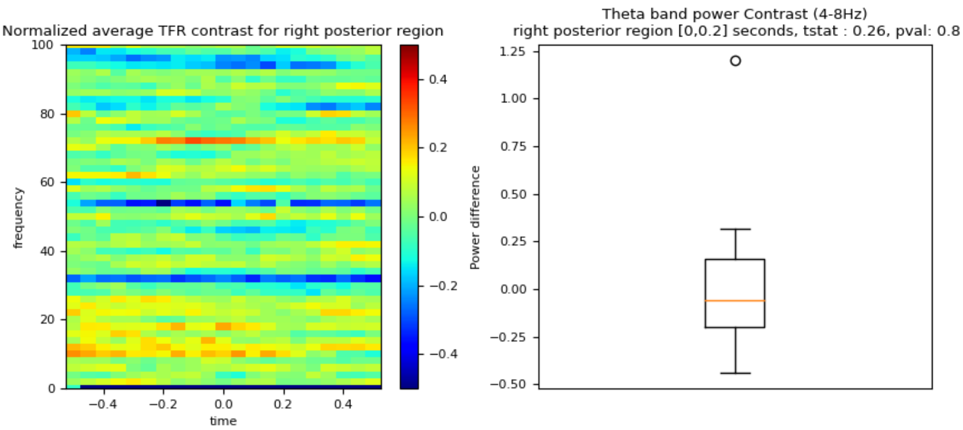

Figure 12: Testing for any statistically significant difference between experimental and control condition in contrast for the right posterior region in the time interval of [0,0.2] seconds.

For the regions of interest as defined in the study, we find no significant difference in activity between the experimental (ambiguous sentences) and control (non-ambiguous sentences) conditions in the Theta band (4-8 Hz) in the [0,0.2] time interval after onset of the critical item for the late learners of German group.

Studies such as the one conducted here where neural activity is analyzed to localize cognitive processes can help shed light into how tasks such as understanding the meaning of a sentence is carried out. For a person fluent in German, understanding a German sentence is an effortless endeavor, given they are not suffering from any cognitive deficits. However, for persons who suffer disabilities due to genetic conditions or physical trauma to the brain, this may not be the case. Understanding the regions of the brain that contribute to cognitive tasks may lead us to better understanding of the mechanisms behind the processes and help develop tools, such as brain-computer interfaces, to assist patients.