Master of Science

# Difference of speech-evoked electroencephalography between native and non-native listeners

The Graduate School

of the University of Ulsan

Department of Biomedical Engineering

Luong Do Anh Quan

# Difference of speech-evoked electroencephalography between native and non-native listeners

Supervisor: Jihwan Woo

A Dissertation

Submitted to

the Graduate School of University of Ulsan

in partial Fulfillment of the Requirements

for the Degree of

Master of Science

by

Luong Do Anh Quan

Department of Biomedical Engineering

University of Ulsan, Korea

June 2021

# Difference of speech-evoked electroencephalography between native and non-native listeners

This certifies that the master's thesis

of Luong Do Anh Quan is approved.

Committee Chair, Dr. Kyo-in Koo

Committee Member, Dr. Jihwan Woo

Committee Member,  Dr. Young-joon Chee

Department of Biomedical Engineering

University of Ulsan, Korea

June 2021

# ACKNOWLEDGMENT

In the beginning, I would like to send my special thanks to Professor Woo Jihwan, who offered me a great chance to be a student as well as a researcher in the CNE lab, which I have earned great experiences and valuable times. I always appreciate his gentility and carefulness for me as well as all his students.

Besides, I also want to express my thanks to my labmates, my Vietnamese friends in Ulsan, and Vietnam, who have always stood by me and supported me a lot during these two years. These helps are considered a great motivation that forces me to study and endlessly improve myself on my career path.

Next, my acknowledgment of my family is perpetual. I gratefully acknowledge my family's support, for their endless encouragement without conditions, base on which I can pursue academic study in the most comfortable way.

Finally, I would like to thank all the thesis defense committee members who helped me review my thesis. I believe I could do better in my research with your valuable comments.

Thank you,

Luong Do Anh Quan

# Contents

# ABSTRACT

Language is one the factors that can affect people's understanding depending on the proficiency of the speaker. The natives comprehend their language easily as compared to non-natives. This may be associated to the difference in brain responses to speech stimuli between native and non-native speakers. Previous studies have proposed sentence comprehension models, such as acoustic-phonetic processor, semantic relations, grammatical relations, prosodic processes and interpretation. This study provides insight into discrepancy in brain activation pattern among speakers (native vs non-native) based on the aforementioned model but only in regards to semantic and grammar. In addition, the cross-correlation function, which is computed from Event Related Potential and speech features, is extracted as a feature of brain activation. Our results show the language effect in early and late processing in both native and non-native speakers. It also suggests that in the case of a different language, our brain tend to concentrate more on syntactic information rather than semantic information.

## List of figures

## List of tables

# 1. INTRODUCTION

## 1.1. Research basis and objective

In daily life, sentence comprehension plays a crucial role in speech communication due to its ability to help us communicate effectively. However, there are some factors that may limit one's comprehension of sentences or speeches, such as hearing impairment, brain injury, speak-in-noise condition and subject condition. In addition, the use of foreign languages could be considered as a language barrier. For instance, if a person listens to two sentences (one in their first language (L1) and one in second language (L2)), they will comprehend the sentence in their first language better compared to the one in the second language (depending on their proficiency in the second language). Therefore, the choice of language may have an influence on human's speech comprehension to some degree.

Recently, Electroencephalogram (EEG) has become a popular and efficient method for monitoring and analyzing the cognitive process; in addition to being noninvasive, inexpensive, fast and having easy recoding procedures, it also has a good temporal resolution. The Event Related Potentials (ERPs) is computed from EEG signal such that it inherits the advantages of EEG. In addition, ERPs improve the ability in term of stimulus locking time to acquire very high temporal resolution signal. Nevertheless, ERPs still have some limitation in regards to difficulty to disentangle multicomponent activity. On the other hand, previous study shows that phonetic and phonemic components can affect adult cross-language perception (Werker & Tees, 1984). Herein, the cross-correlation method which reflects the correlation between ERP and speech features of stimuli was utilized.

Previous studies have shown the effects of language in listener by using ERP (Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Friederici, 2011; Kutas & Hillyard, 1980, 1983; Rauschecker & Scott, 2009). These studies discovered how brain process information in speech such as acoustic information (Rauschecker & Scott, 2009), semantic relations (Kutas & Hillyard, 1980), grammatical relations (Kutas & Hillyard, 1983), integrated information (Bornkessel et al., 2005). Then, the language comprehension was sketched to present the speech perception process (Friederici, 2011). However, most of studies were conducted in native speakers and just a few in non-native speakers. Then, (Correia, Jansma, Hausfeld, Kikkert, & Bonte, 2015) found the relationship between words and language invariant semantic-conceptual representations in spoken words. Although this study showed the neural mechanisms underlying conceptual encoding in comprehension and production, it was just evaluated at word stimuli level not in sentence stimuli level. This inspired the upcoming studies in investigating the effect of speech perception in non-native speakers using continuous speech stimuli.

Besides, oscillatory response to speech plays a vital role in sentence's comprehension process (Bastiaansen, Magyari, & Hagoort, 2010; Davidson & Indefrey, 2007; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Kielar, Panamsky, Links, & Meltzer, 2015; Klimesch, Doppelmayr, & Hanslmayr, 2006; Krause, 1999; Wang et al., 2012; Willems, Oostenveld, & Hagoort, 2008). These studies showed the working mechanism of frequency bands in different cases by analyzing event-related synchronization (ERS) and the event-related desynchronization (ERD), which reflect the increase and decrease in power, respectively. However, there is a lack of work investigating the difference in oscillatory dynamics between native and non-native listeners. Though different frequency bands, including theta (4-7 Hz), alpha (8-12 Hz), and beta (15-30 Hz), undertake different cognitive tasks, this study will focus on alpha band oscillations because it relates to semantic processes and linguistic (Kielar et al., 2014; Kielar et al., 2015; Klimesch et al., 2006; Willems et al., 2008).

This study aims to understand the cortical mechanism of language processing between native and non-native speakers by measuring the continuous speech EEG signal and using the cross-correlation function calculated based on the speech features of stimuli. This study was conducted by measuring EEG in native (English) and non-native (Native-Korean) speakers according to the stimulation of English sentences. In addition to the clean speech presentation analysis, we want to see the difference between native and non-native speakers' brain activation according to the degraded continuous English speech stimuli. We hypothesize that non-native speaker have stronger language component effects due to difficulty in understanding speech in foreign language.

1.2. Basic brain anatomy and function

The major motor and sensory regions of the cerebral cortex:

**Table 1**: Location and function of brain areas

| Lobe | Region | Function |
|---|---|---|
| Primary motor cortex | Frontal lobe, form the anterior border of the central sulcus | Voluntary control skeletal muscle movement if it is stimulated (By controlling pyramid cells in the brain stem and spinal cord) |
| Primary sensory cortex | Parietal lobe, form the posterior border of the central sulcus | Conscious perception of somatic sensory information from the receptors for touch, pressure, vibration, taste or temperature |
| Visual cortex | Occipital lobe | Conscious perception of visual information |
| Auditory cortex | Temporal lobe | Conscious perception of auditory information |
| Olfactory cortex | Temporal lobe | Conscious perception of olfactory information |
| Gustatory cortex | Anterior portion of the insula and adjacent portions of the frontal lobe | Conscious perception of stimuli from taste receptors of the tongue and pharynx |
| Association areas | All lobes | Integrate and process sensory data, process and initiate motor activities |

**Association Areas**

*Sensory association areas* are cortical regions that monitor and interpret the information arriving at the sensory areas of the cortex. Each has its own association areas. For examples:

- ***Somatic sensory association area*** monitors activity in the primary sensory cortex.
- ***Visual association area*** monitors patterns of activity in the visual cortex and interpret the results.
- ***Auditory association area*** monitors sensory activity in the auditory cortex.

*The somatic motor association area* or *premotor cortex* takes responsibility for the coordination of learned movements. The neuron in the primary motor cortex must be stimulated by neuron in other parts of cerebrum.

**Integrative Centers**

*Integrative centers* are areas that receive information from many association areas and direct extremely complex motor activity. Integrative centers concerned with the performance of complex processes are restricted to either the left or the right hemisphere. These centers include the general interpretive area and the speech center.

**Table 2**: Location and function of integrative centers

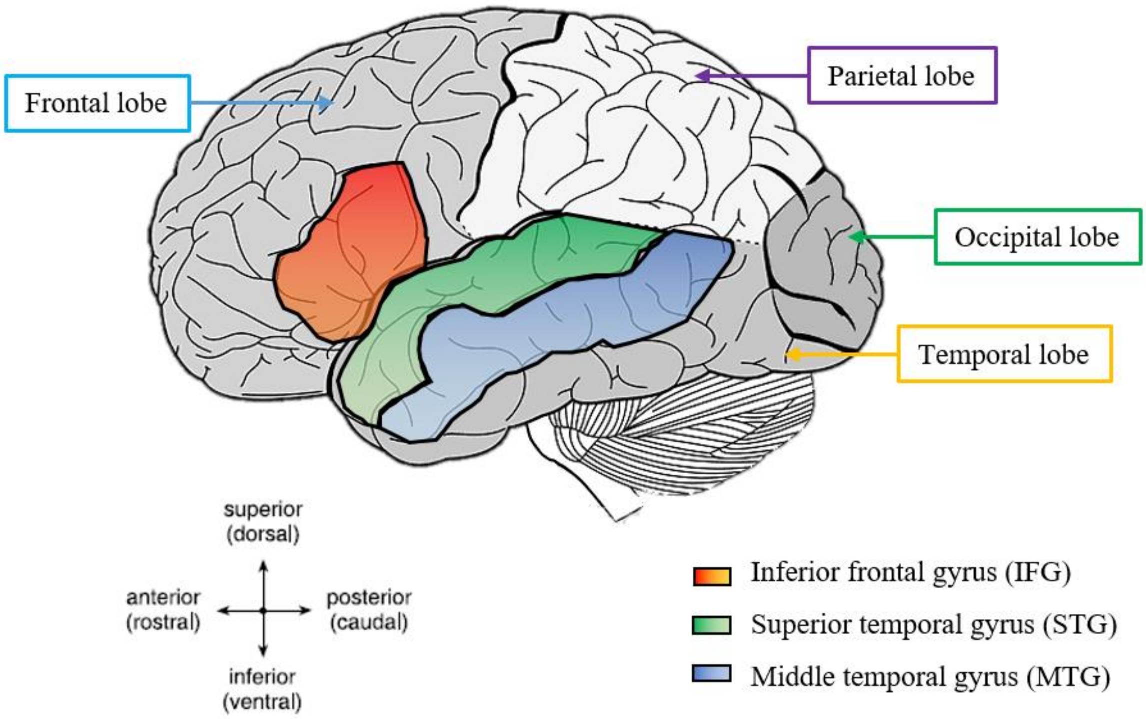
| Area | Input | Location | Function |
|---|---|---|---|
| **The General Interpretive Area** or **Wernicke area,** also called **the gnostic area** receives | Receive information from all the sensory association areas | In only one hemisphere (typically left) | Play an essential role in personality by integrating sensory information and coordinating access to complex visual and auditory memories |

| | | | |
|---|---|---|---|
| **The Speech Center.** Some of the neurons in the general interpretive area innervate the **speech center** also called **Broca area** or **the motor speech area** | Receive feedback from the auditory association area | Lie along the edge of the premotor cortex in the same hemisphere as the general interpretive area (left) | Regulate the patterns of breathing and vocalization needed for normal speech. |
| **The Prefrontal Cortex** | Receive information from the association areas of the entire cortex | The prefrontal cortex of the frontal lobe | Perform abstract intellectual functions as predicting the consequences of events or actions. |
| **Brodmann Areas** | | | Some correspond to known functional area |

1.3. The language-related circuit

From different overview (Friederici, 2002; Hickok & Poeppel, 2007; Vigneau et al., 2006), the language-relevant cortex includes Broca's area in the inferior frontal gyrus (IFG), Wernicke's area in the superior temporal gyrus (STG), as well as parts of the middle temporal gyrus (MTG) and the inferior parietal and angular gyrus in the parietal lobe were involved in language perception (Figure 1). Each area plays their specific roles and connects with others to achieve speech understanding.

**Figure 1**: Anatomical and cytoarchitectonic details of the left hemisphere.




1.3.1 Role of temporal lobe

In general, the temporal includes two parts: the anterior temporal cortex and the posterior temporal cortex. The anterior temporal cortex was claimed to be involved in semantic and syntactic processes and primarily combinatorial in nature. The posterior temporal lobe is involved in language processing, and its function appears to be primarily to integrate different types of information. For sentence processing, this might mean the integration of semantic and syntactic information.

### 1.3.2 Role of IFG

The IFG, in particular Broca's Area, has long been known to support language production though its function is still a matter of considerable debate. These overview (Bookheimer, 2002; Friederici, 2002; Grodzinsky, 2000; Grodzinsky & Friederici, 2006) show evidences that some factor such as syntactic complexity, a combination between syntactic complexity and working memory or a combination between syntactic complexity and task demands, may affect the neural response in the IFG.

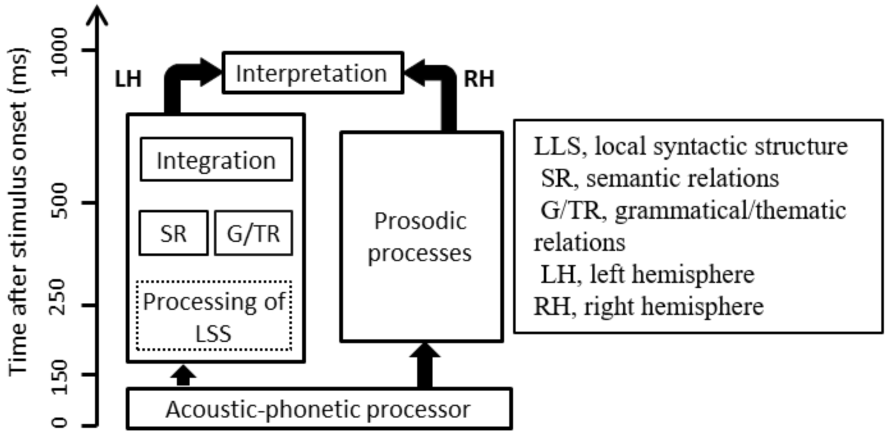
### 1.3.3 The auditory language comprehension model

Figure 2 describe the time course of language perception process. As the figure shows, the process was divided into small subtasks, which is executed in different cortices and in different time courses. The N100 reflecting the process of acoustic-phonetic is activated in the early time (0 – 100 ms) at primary audio cortex. The information extracted from the first step is then separated to many threads toward to both left hemisphere and right hemisphere. The left hemisphere is generally known to be responsible for processing syntactic and semantic factor, while the right hemisphere analyzes the prosodic information. Then, these features will be interrelated as the last step.

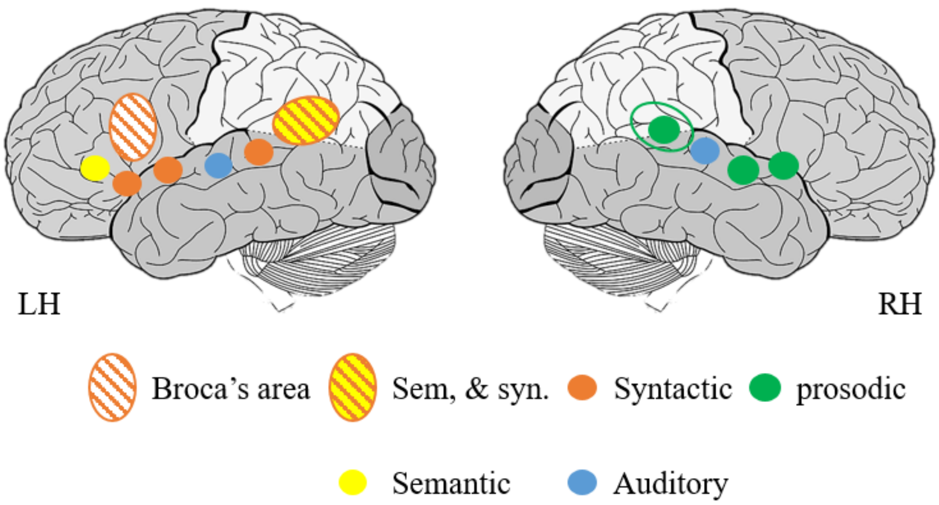
In the language component processing step, which occur in left hemisphere between 150 ms and 800 ms, there are 3 major periods. The first period (150 – 250 ms) is responsible for processing phrase structure. This period is normally presented by the early negativity. The second period (250 – 500 ms) is responsible for sematic and grammatical relations, presented by the N400 and anterior negativity, respectively. Last period is from 500 ms to 800 ms. This period integrates information from previous periods and presented by the late positivity.

**Figure 2:** Model of auditory sentence comprehension.

**A. Language comprehension model**



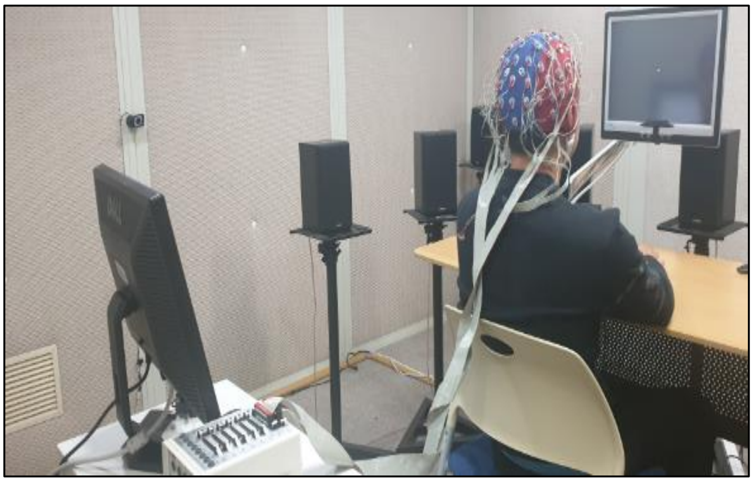

**B.** The brain basis of auditory language comprehension



## 1.4. Electroencephalography

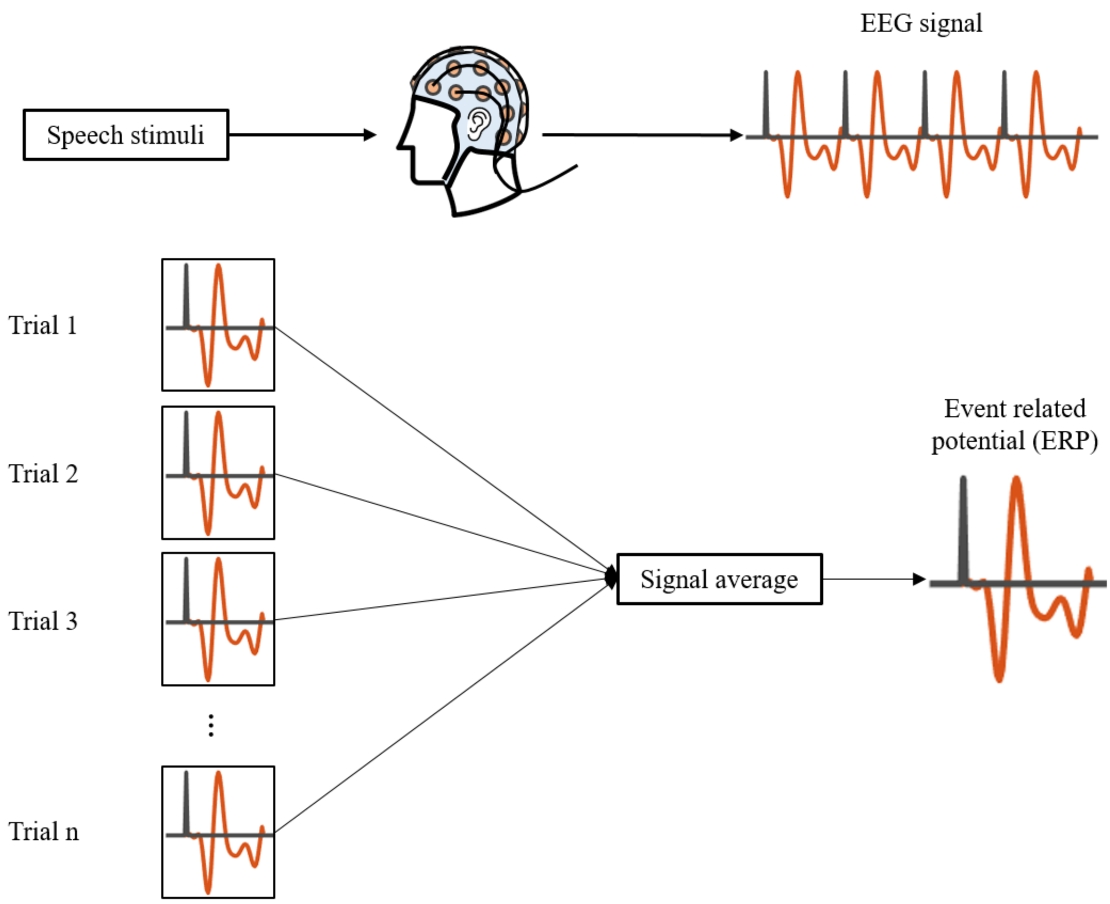
Electroencephalography (EEG), which is a measured method recording the electrical activity of the brain, was first recorded on human by a German physiologist, Hans Berger in 1924. Nowadays, it is widely used to detect abnormalities in your brain waves or diagnose other disorders that influence brain activity, such as Alzheimer's disease, certain psychoses, and a sleep disorder. This is because EEG’s advantages, which are a good time resolution, a non-invasive method, a good method for state effects and fast, inexpensive and easy recording procedures. However, EEG has a low spatial resolution because it captures the brain signals through only electrodes employed over our scalp (Number of electrodes may be 16, 32, 64 or 128 depend on the system). Figure 3 shows an instance of EEG scalp and EEG signals. EEG also has a low signal-to-noise ratio.

**Figure 3**: EEG measurement



Recently, many studies took advantages of EEG to investigate our brain perception mechanisms in many aspects: auditory perception, visual perception, motor action and so on. The event-related information is extracted by using ERPs, which is achieved by averaging out the EEG signals. By averaging the EEG signals, we can remove the random EEG and obtain the event-related response because these signals are phase locked and time locked together. The reason for this time-locking and phase-locking is that the neuron spikes will be excited by the stimuli. Back to the ERPs, it includes a series of peaks and troughs, which are known as ERPs’ components. Components are usually named as a combination of a letter “P” (positive) or “N” (negative) and a number indicated its time point, for example, N100 is a negative component occurred at 100ms after stimulus onset. Figure 4 shows the example of ERPs in auditory task.

**Figure 4**: Example of ERPs components and how to achieve ERPs



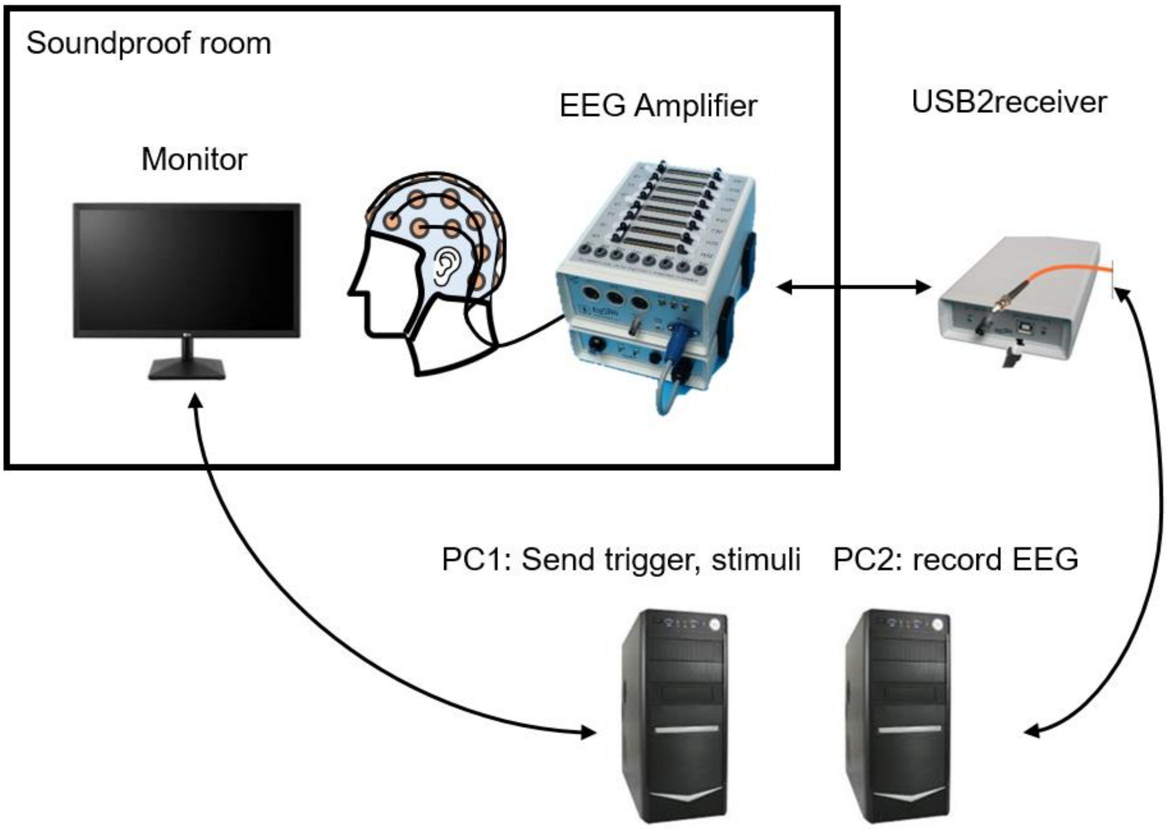
# 2. BASIS THEORY

### 2.1. EEG recording system

#### 2.1.1. Experiment diagram

Thirteen native (American) and thirteen non-native (Korean) normal hearing participants (male: 10, female: 17) were recruited in this study. Particularly, native speakers were recruited in University of Iowa, USA and non-native speakers were recruited in University of Ulsan, Korea. The data of one male, a non-native participant, was removed due to his performance during the experiment. English was reported to be the first language for all the native subjects (L1). In case of non-native speaker, Korean was reported to be the first language and they have a low-level English skill (L2). All subjects sat on a chair at a distance of 1 m from the loudspeaker in a sound-proof room and watched silent movies during the experiments. All study procedures were reviewed and approved by the two Institutional Review Boards of the University of Iowa and the University of Ulsan. Figure 5 describe the diagram of our system.

**Figure 5:** EEG recording diagram



Devices and toolboxes

- ❖ EEG signal was recording using Biosemi system which includes:
  - Active Electrodes and Headcaps
  - AD-box
  - USB2 Receiver
  - Charger
  - Battery box
  - Analog Input Box
  - ActiView software (for monitoring EEG signal)

❖ Auditory stimuli were manipulated using Psychotoolbox and Matlab environment. The Psychotoolbox embedded in Matlab provides the synchronization for sending sound to speakers and sending trigger to the Biosemi system, which help to extract the stimuli onset during processing the EEG signals. Then, the Biosemi system record the signals and save to the server for offline processing.

### 2.1.2. Experiment procedure

In preparation for the experiment, participants were asked to wear the electrode holders cap which holders. After that, electrode holders were filled with electrode gel with a syringe to improve the contact between the electrodes and our scalp; then the electrodes were attached on the cap and connected to Biosemi system. All of the measuring procedure was monitored on PC1 (Figure 5) to track the subject’s state.

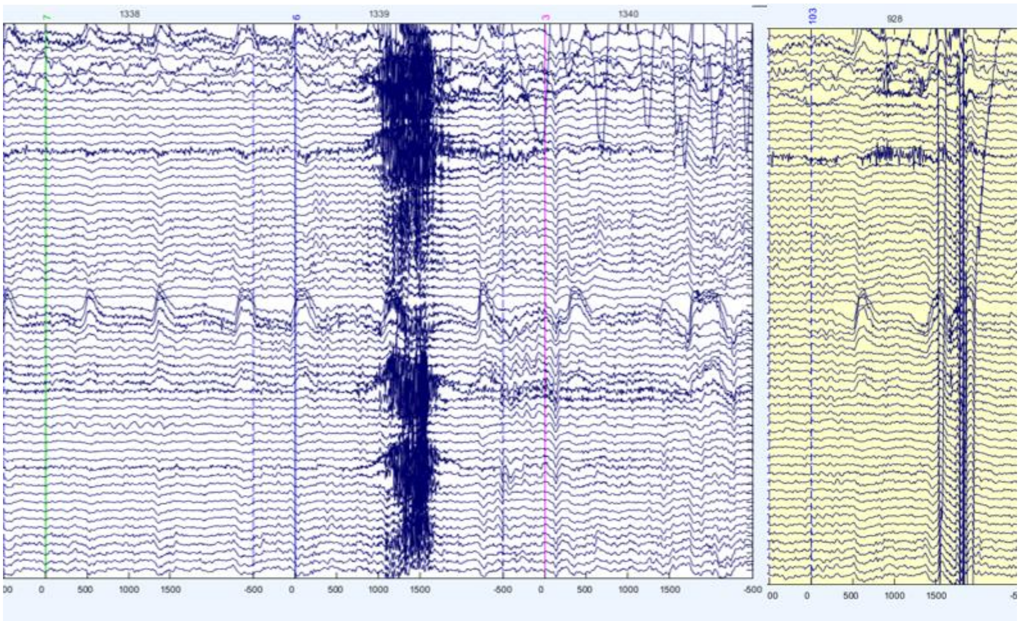
The experiment consisted of two kinds of stimulus: natural and vocoded sentences. At first, ten continuous English sentences from a revised-speech which underwent noise test were selected (R-SPIN reference). Then, ten eight-channel vocoded sentences corresponding to ten natural sentences were made via a CIS strategy (CIS reference). Each sentence was less than 2.5 seconds in length. Subjects were exposed to both natural and vocoded sentences which were played randomly through the loudspeaker in the passive listening task (vocoded sentences were played prior to natural sentences in order to reduce the learning effect from the natural condition). The stimuli contained 10 natural and 10 vocoded sentences and each sentence was repeated 100 times. The time interval between the trials was 3 seconds. Electroencephalograms (EEG) were recorded while the subjects were undertaking the listening tasks. Subjects were asked to maintain the same body posture in order to minimize artifacts.

2.2. Data processing procedure

2.2.1. EEG recording and processing

The EEG signal was recorded by 64 Ag-AgCl active electrodes with their locations done according to the location of international 10-20 system electrode placement at a sampling rate of 2048 Hz. The signal was amplified using a Biosemi Active 2 system and filtered with a bandpass filter at 2-57 Hz. The raw EEG signal was then re-referenced and down sampled to 256 Hz. The down sampled signal was epoched from 500 ms to 2500 ms before and after a stimuli onset, respectively. An inspection was done to determine the defective trials which were affected by either noise or the subject's fatigue.

**Figure 6:** Example of bad epochs (noise or subject's scratching)



Normally, the EEG signals are easily effected by artifact such as eye movements (EOG), eye blinks, EMG or noises. These artifacts are not related to the stimuli, but they still have a big amplitude that may ruin the signals. Therefore, Independent Component Analysis (ICA) was applied to detect and remove the artifacts. ICA decomposes signals to selected number of components with their power density spectrum, distribution and time series values so we can make a judgement which one is artifact. Figure 7 and 8 indicate the example of eye blink artifact and the impact of removing this artifact, respectively.

**Figure 7**: Example of eye blink artifact

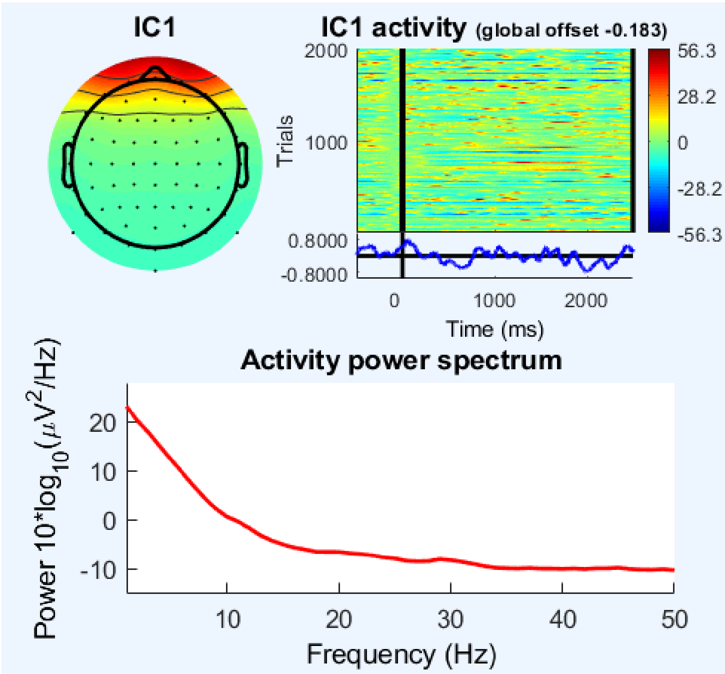
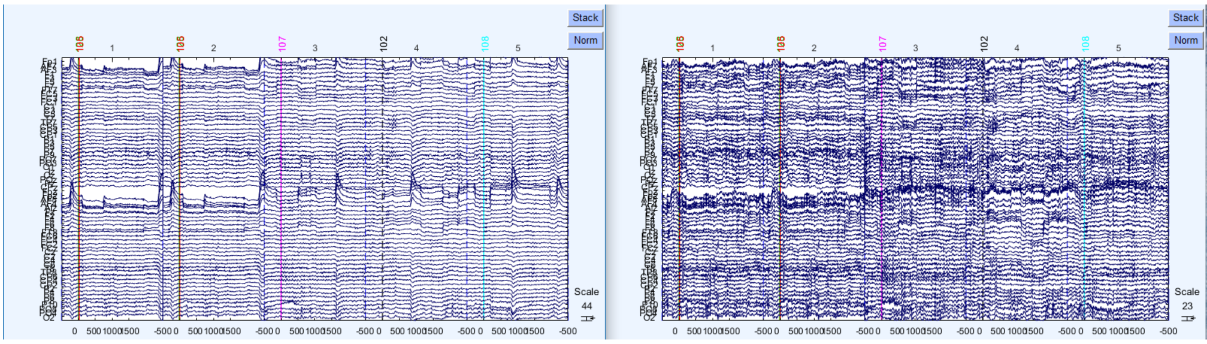



**Figure 8**: The example of removing eye blink artifact from the signals



After removing artifacts, epoches that exceeded the range -100 uV to 100 uV were excluded from the data. In summarize, EEG signals recorded during the experiments were preprocessed following below steps:

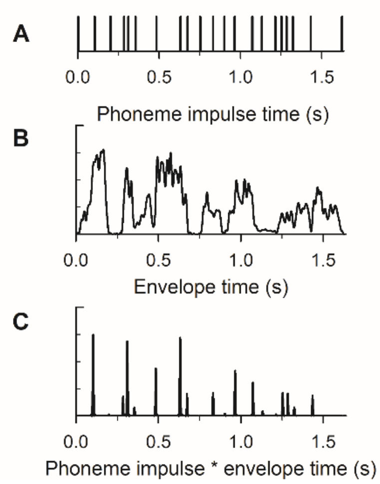
- Re-reference
- Filter (0.5-57, order: 5$^{th}$)
- Resampling to 256 Hz
- Independent components analysis
- Remove artifacts

- Epoch based on trigger (-0.5-2.5 s)
- Remove out of bounds epochs (-100-100 μV)

2.2.2.EEG analysis

The ERPs were calculated by averaging the remaining epochs of each sentence after eliminating out-of-bound epochs and subtracting the baseline, which ranged from -200 to 0 ms. Finally, the ERPs were filtered through a bandpass filter of 1-15 Hz. The phoneme onset impulse train which is actually the onset times of phonemes in each sentence was extracted for natural and vocoded sentences using Webmaus. Figure 9A shows an example of the phoneme onset impulse train for the sentence "Bill heard Tom called (enquire?) about the coach". Figure 9B and 9C show the speech envelope and the combination between speech envelope and phoneme impulse train, respectively. Then, the neural response to continuous speech, illustrated as in figure 1E, was calculated based on a cross-correlation function between ERP (figure 10) and the combined features corresponding to each sentence. The cross-correlation function having a time-lag of -200 to 800 ms was obtained by applying 5$^{th}$ Butterworth high-pass filter with a cut-off frequency of 1 Hz. Finally, the grand average was computed across all subjects for each case involving native natural sentence, native vocoded sentence, non-native natural sentence and non-native vocoded sentence.

**Figure 9**: An example of (A) phoneme impulse train for whole-sentence, (B) speech envelope, (C) The combination of phonetic information and speech envelope.

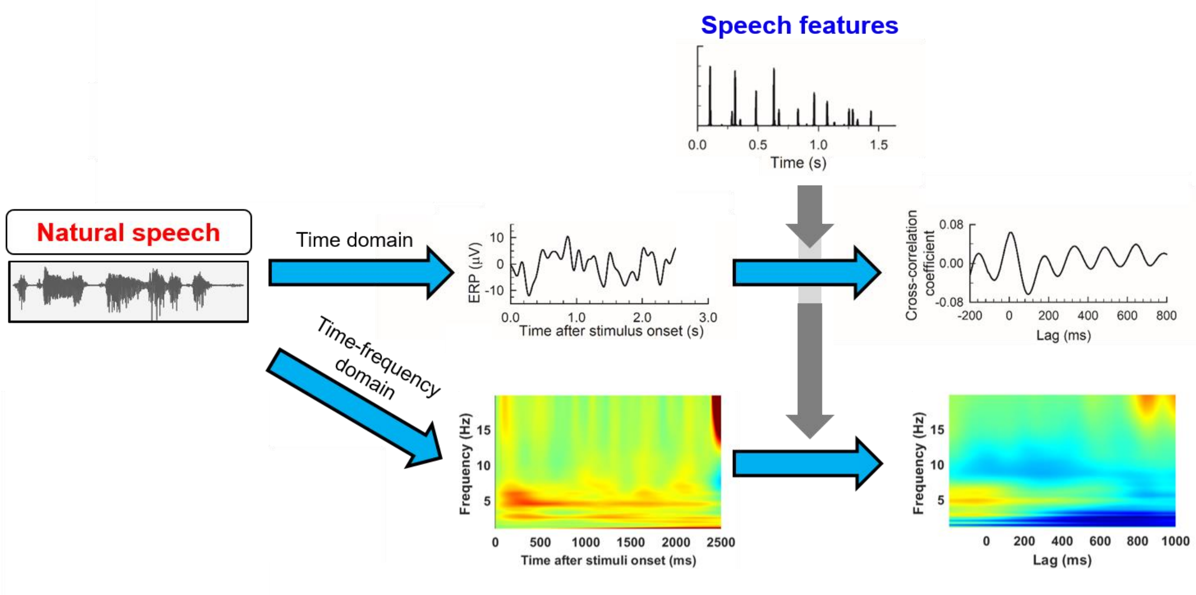


For spectral analysis, each epoch's time-frequency representation was decomposed by using a Morlet wavelet transform. The event-related spectral perturbation (ERSP) of each sentence was achieved by averaging all spectrum of corresponding trials. The percentage change normalization, which is computed by the following formula:

$$prctchange_{tf} = 100\,\frac{activity_{tf} - \overline{baseline_f}}{\overline{baseline_f}},$$

was then applied for ERSPs to indicate the power changes during speech perception. Figure 10 present the ERSP spectral and the power modification at 10 Hz. To observe how oscillatory response constraint to stimuli, the cross-correlation function between ERSP and the stimuli was assessed as described in Fig. 10. In the last step, the grand average cross-correlation spectrum was computed across all subjects for each case, as in CSEP analysis.

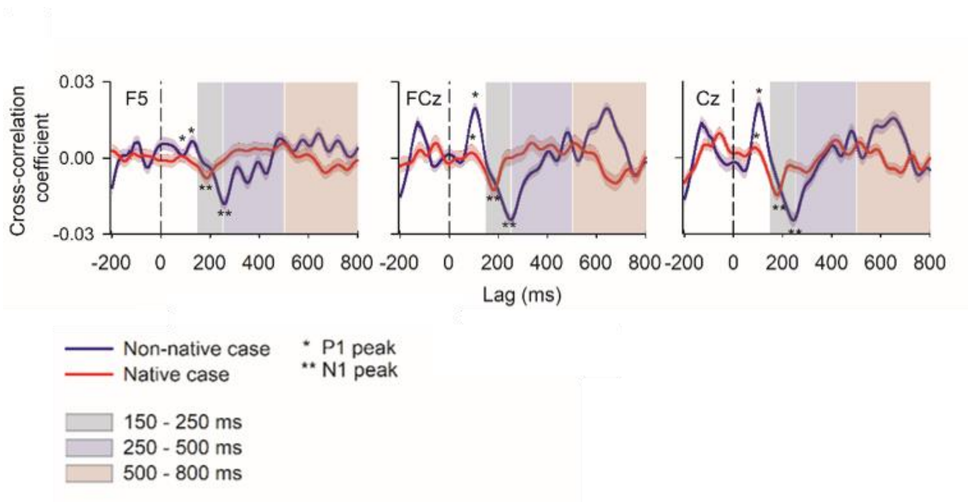
**Figure 10**: Method for analyzing neural response to continuous speech stimuli



# 3. RESULTS

The result was analyzed based on the model of auditory sentence comprehension, which is illustrated in Fig. 2A (Friederici, 2011). However, the components involved in syntactic analysis and semantic processing from 150 to 800 ms are focused. Figure 11 presents the grand average cross-correlation coefficients of several channels and descriptions of 4 major effects: P1 delay, N1 delay, positive trend, and negative trend in response to 3 kinds of stimuli. From the plot of F5, Cz, and FCz channels, it can be observed that the non-native listeners' P1-N1 peaks were delayed about 100 and 50 ms compared to those of native listeners (Figure 11). The graph also suggests that the non-native case's amplitude tends to be more negative in the period from 250 to 500 ms and more positive after 500 ms compared to that of the native case. Notably, F5 and FCz channels have similar shape of brain activation as Cz channel.

**Figure 11:** The grand average cross-correlation coefficients at F5, FCz, and Cz channels in natural sound case. In each plot, P1 (first positive peak) and N1 (first negative peak) were denoted using * and **, respectively. The time courses of 150 – 250, 250 – 500 and 500 – 800 ms were illustrated by using gray, purple and orange areas, accordingly. Standard deviation of cross-correlation coefficients across subjects were presented by the shaded areas.



To compare the mean cross-correlation coefficient in 3 main syntactic and semantic processing steps between native and non-native listeners, the mean value of the grand average cross-correlation coefficients for 3-time ranges (150 – 250 ms, 250 – 500 ms, and 500 – 800 ms) were computed for all the 64 channels. The cross-correlation coefficients at all sampled time were used to calculate the mean value in each time range. The mean value in the non-native case is subtracted by that in the native case to achieve a comparison.

**Figure 12:** Topographies of mean cross-correlation coefficient between ERP and stimuli in each auditory language processing phase and the differences in that between native and non-native natural sound case. The mean cross-correlation coefficient was calculated for each case by averaging all cross-correlation values in corresponding time range for each electrode. The difference was evaluated in 3 main processing phases (150 – 250, 250 – 500 and 500 – 800 ms).

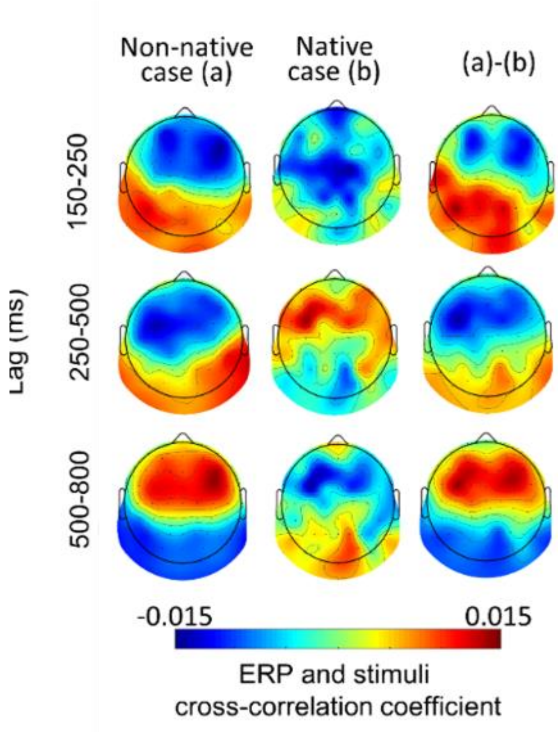



Fig. 12 illustrates anterior negativity in both native and non-native speakers in the first language processing period. However, it can be seen that non-native speaker has stronger negativity effect. In Fig. 11, the N1 peaks in the left frontal and central regions (F5, FCz, and Cz) are observed in the non-native case at roughly 200 ms after the sentence onset. From 250 to 500 ms, the non-native case's power is more negative than that of the native case in both the anterior cortex and the central site. The 250 – 500 ms time range also contains a negative peak in the non-native case at about 290 ms after trial onset in the left frontal site (F5), central frontal site (Fz), and central site (Cz) (Fig. 12A). Figure 11 shows a positive peak of the non-native case in frontal-central sites (Fz, FCz, and Cz) at approximately 660 ms. In other words, the later phase (500 – 800 ms) presents a positive trend in the frontal-central in the case of non-natives.

**Figure 13:** Difference between non-native's grand average spectrum and native's grand average spectrum (non-native's – native's) at F5 in natural sound case and difference topographies in different periods (The time interval is between -500 ms and 2500 ms relative to the onset of the continuous speech stimuli, with a time step of 500 ms.). (The signal on the top presents the speech sentence in experiment with the duration less than 2 seconds)

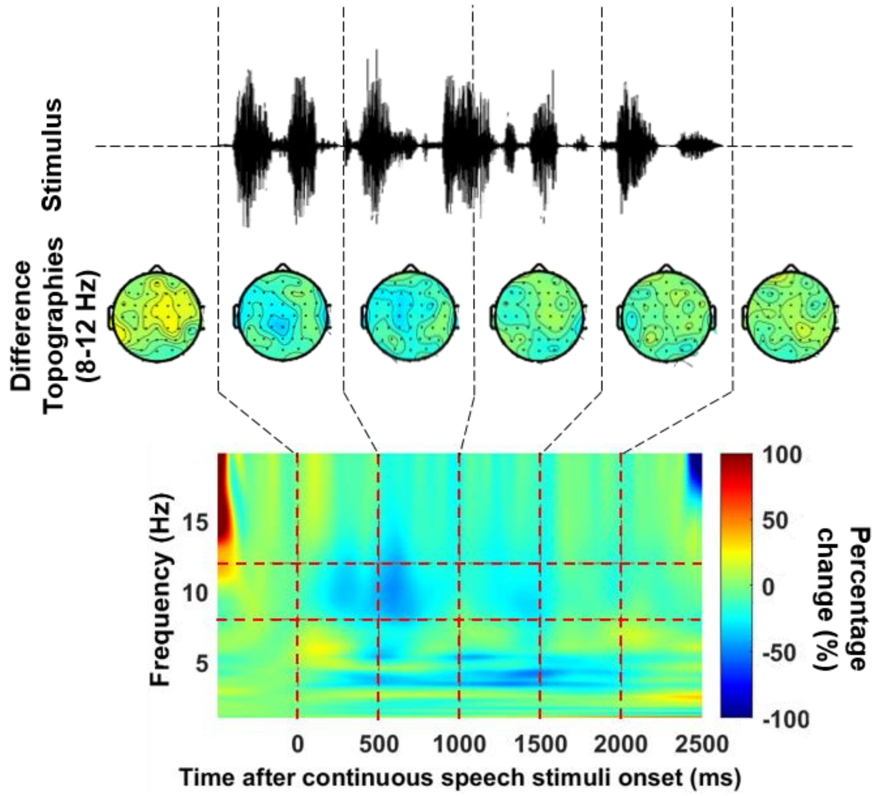
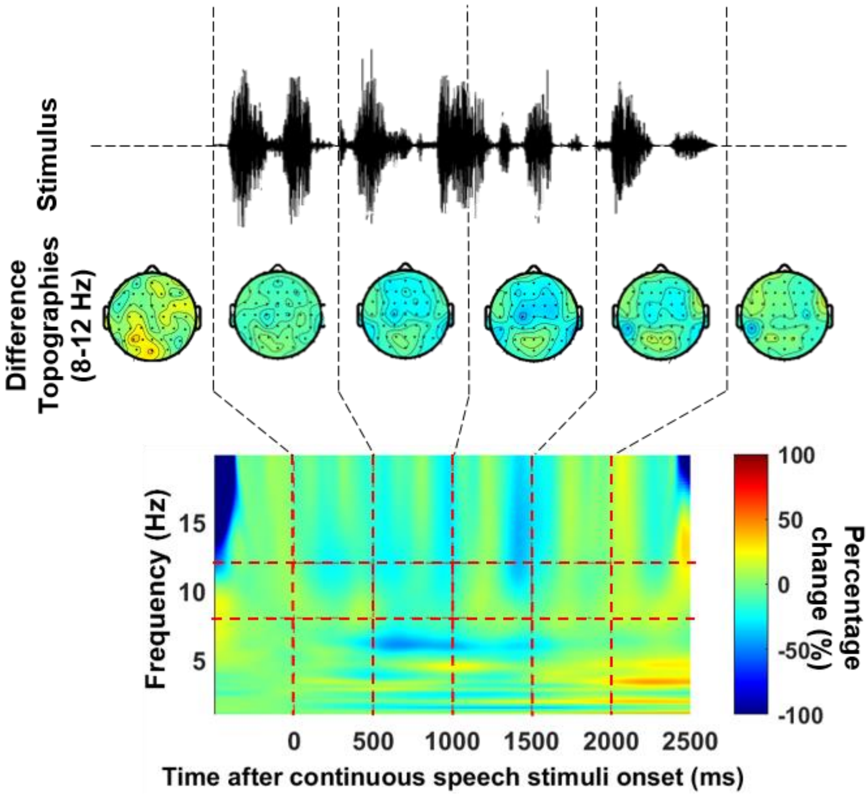


**Figure 14:** Difference between non-native's grand average spectrum and native's grand average spectrum (non-native's – native's) at Fz in vocoded sound case and difference topographies in different periods (The time interval is between -500 ms and 2500 ms relative to the onset of the continuous speech stimuli, with a time step of 500 ms.). (The signal on the top presents the speech sentence in experiment with the duration less than 2 seconds)




Figures 13 and 14 show the difference spectrum and the alpha-band difference topographies between non-native's and native's oscillatory responses in natural and vocoded sound conditions. It can be seen that the natural sound condition presents non-native listeners' ERD in the left frontal, left temporal, central, and parietal cortices during speech duration. On the other hand, the vocoded sound condition reveals the non-native listeners' ERD in the central and frontal sites more than in the left frontal and left temporal sites though this ERD might not result from the speech.

**Figure 15:** Difference between non-native's cross-correlation spectrum and native's correlation spectrum (non-native's – native's) at F5 in natural sound case and alpha band power topographies in 4 time periods. The significate difference areas (p value < 0.05), which are analyzed by non-parametric permutation test, is located inside the black boundary.

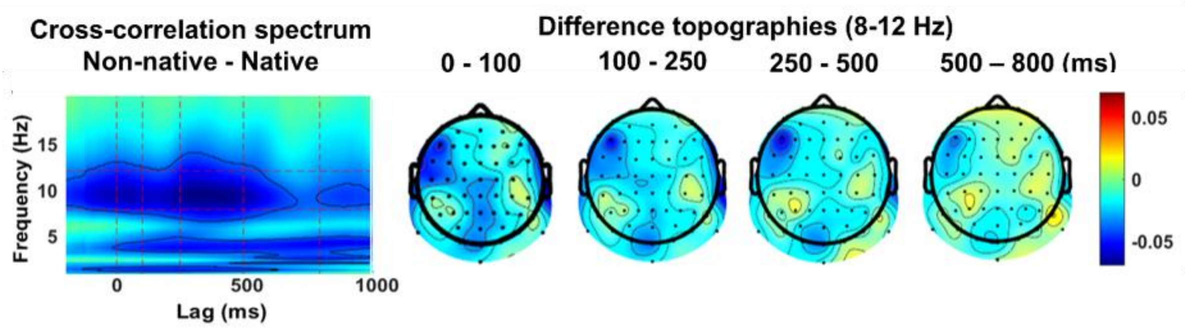
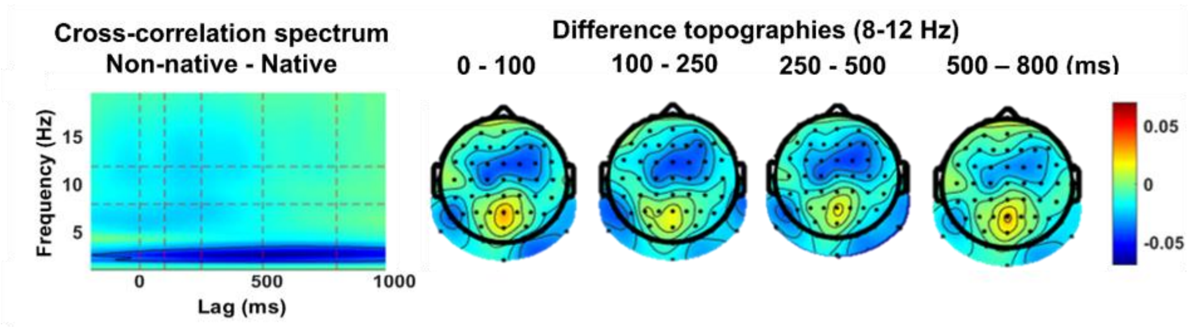



**Figure 16:** Difference between non-native's cross-correlation spectrum and native's correlation spectrum (non-native's – native's) at Fz in vocoded sound case and alpha band power topographies in 4 time periods. The significate difference areas (p value < 0.05), which are analyzed by non-parametric permutation test, is located inside the black boundary.




To investigate how brain oscillation constraint to speech, this study examines the difference cross-correlation spectrum and the corresponding alpha-band difference topographies in both natural and vocoded sound condition. The difference graphs show the significant alpha ERD of non-native listeners (p values<0.05) in the left frontal, left temporal, central, parietal, and occipital areas (Fig. 15). The alpha ERD in the left frontal remain quite strong until the later phase, while ERD in the remaining regions decrease time by time. The alpha ERDs in response to speech stimuli show more substantial effect in the left temporal, central, and occipital sites (Figure 15). Figure 16 indicates the difference in oscillatory response for vocoded sound condition. Even when topographies show some negative values in the range of alpha band (8-12 Hz), the cross-correlation spectrums do not show significant differences (p values<0.05) in any stimuli case.

# 4. DISCUSSION AND CONCLUSION

In this study, the delay in P1 and N1 peak between the native and non-native case was apparent. The distinction in brain activation patterns across American and Korean persons when listening to English sentences is also presented in three main stages of the auditory language processing model (construct a local syntactic structure, process semantic and syntactic relations, and integrate this information).

According to auditory sentence comprehension, early left anterior negativity (ELAN) is the first sentence-level component responsible for analyzing the information on the word category (verb, preposition, etc.). The initial local phrase structure can then be achieved by extracted these information (Friederici, Pfeifer, & Hahne, 1993; Friederici & Weissenborn, 2007; Isel, Hahne, Maess, & Friederici, 2007; Kubota, Ferrari, & Roberts, 2003; Neville, Nicol, Barss, Forster, & Garrett, 1991). Here, ELAN is evoked in the native case, and this early negativity may be a result of the ease of categorizing word information (function word, morphological information, etc.) from the stimulus sentence (Dikker, Rabagliati, & Pylkkanen, 2009; Neville et al., 1991). In addition, non-native's topographies also indicate the ELAN but stronger. This coincidence may relate to the change in P1 and N1 peak in response to different stimuli. Table 3 shows several significant differences in the N1 peak and N1 latency in neural response to speech feature. This observation is consistent with previous studies regarding different onsets and topography patterns due to syntactic anomalies, including syntactic violations or morphosyntactic violations (Friederici & Frisch, 2000; Gunter, Friederici, & Schriefers, 2000; Hahne & Jescheniak, 2001; Hinojosa, Martín-Loeches, Casado, Muñoz, & Rubia, 2003).

**Table 3:** Mean and Standard Deviation (Std) of N1 peak latency, amplitude of 3 channels in 3 stimuli cáe were calculated over all native and non-native subjects (Location was identified based on 10-20 system electrode placement). Difference was analyzed by using Wilcoxon signed rank test

| Stimuli | Ch | Location | Time after stimulus onset (ms) | | | Amplitude | | |
|---|---|---|---|---|---|---|---|---|
| | | | Non-native Mean (Std) | Native Mean (Std) | p | Non-native Mean (Std) | Native Mean (Std) | p |
| Combined feature | F5 | Left frontal | 266 (49) | 218 (66) | * | -0.022 (0.008) | -0.011 (0.011) | *** |
| | FCz | Post frontal | 257 (43) | 201 (51) | | -0.029 (0.010) | -0.016 (0.010) | *** |
| | Cz | Central | 240 (30) | 193 (44) | ** | -0.030 (0.011) | -0.018 (0.015) | ** |

** p < 0.05; *** p < 0.01; * p < 0.10; Ch: Channel; Std: Standard deviation

Regarding the second component of language, which consists of (left) anterior negativity (LAN) and N400, results exhibited a negative trend in the frontal site (F5 and Fz) and central site (Cz), which are typical regions for AN and N400 components, respectively. Nevertheless, the negative power in the frontal area is slightly more than in the central area. Although several studies highlighted that a foreign language's effect evokes the N400 component without mentioning the AN component (Aparicio et al., 2012; Soskey, Holcomb, & Midgley, 2016), this argument could be explained by the level of speech perception (sentence-level versus word-level). It is consistent with the finding that LAN's appearance is a distinct effect of listening to a second language (Friederici & Weissenborn, 2007). Interestingly, in Fig 11, the N1 peak of non-native listeners was delayed to the middle phase compared to the N1 peak of native listeners. The result herein suggested that the latency of the N1 peak may cause the anterior negativity effect. The recent claim is reasonable since both the AN and ELAN correlate to syntactic processing. Therefore, the slowness to build the local phrase structure could lead to some challenges in processing grammatical/thematic relations. Besides, (Knosche, Maess, & Friederici, 1999) revealed the activation of anterior negativity (AN) instead of left anterior negativity (LAN) in the processing of syntactic information, which suggests that the right hemisphere is involved in auditory language processing to support the left hemisphere handles the other tasks.

Previous studies on the comparison between the reanalysis of P600 and repair of P600 reveal distinct distribution patterns in regards to the integration processing step. While the former component showed the frontocentral distribution, the latter's distribution scattered in the centro-parietal site (Friederici, Hahne, & Saddy, 2002). This conclusion supported our result of the frontocentral site's responses across native and non-native cases.

In natural sound condition, the alpha band's topographies in ERSP analysis showed the ERD of non-native case from the speech onset until the end of the speech. The finding proves that this response may result from non-native's speech processing. To be in more specific, the cross-correlation analysis shows the significant ERD in language-related circuits (left frontal, left temporal and central cortices). The alpha ERD distribution is consistent with previous studies about alpha oscillation in sentence (Bastiaansen et al., 2010; Davidson & Indefrey, 2007; Kielar et al., 2014; Kielar et al., 2015; Wang et al., 2012). These studies reported the alpha power decrease in mentioned language-related cortices is induced by the syntactic and the phrase structure violation. This finding supports the hypothesis of syntactic violation, which is illustrated as the ELAN and LAN in CSEP analysis. Besides, the modification of brain activation at alpha band, which declines the power from early-stage (0 – 250 ms) to late-stage (500 – 800 ms), is similar to the power decrease presented in work about word categorize violation (Bastiaansen et al., 2010). The vocoded sound condition's results did not present significant differences (p value<0.05) between non-native and native spectrums at alpha frequency band (Fig.16).

To compare with previous research, which investigate the EEG decoding of spoken words in bilinguals (Correia et al., 2015), this study suggest a novel method to track the neural response to stimuli. While the previous research investigates at word-level stimuli, this study approaches at sentence-level stimuli, which is more natural in daily communication. To the results, at the word-level stimuli, non-natives perform semantic process which transform word from foreign language to first language. However, at the sentence-level stimuli, non-natives showed the more syntactic or grammatical process rather than semantic process.

Altogether, this research has proved the ability of using continuous speech stimuli in investigate language perception in human brain. By using continuous stimuli (e.g. speech, story, music), we can figure out how brain can perceive and process information in the more natural stimuli. Moreover, the result also suggests typical regions, periods, and frequency band for processing language in human brain. These findings could be used as the features for future researches in word or sentence decoding, speech intelligibility prediction, …

The study results show the cross-correlation method's potential to investigate the oscillatory response to continuous stimuli. However, there are some limitations. Firstly, only one speech feature was considered in this case. Therefore, the effect of other features (phonetic features, semantic feature, …) should be discovered in future researches. Secondly, the cross-correlation method's limitation in analyzing two signals with common trends leads to spurious responses in the spectrum (Figure 15 and 16). So, the research about neural response to different stimuli and de-trending signals is required in future work.

In conclusion, this study examines the difference in speech comprehension tasks between non-native and native listeners by apply the cross-correlation function to ERPs and ERSPs. The results suggest that newly learned language speakers show components including ELAN, LAN, and reanalysis P600 in CSEP analysis and alpha ERD in spectral analysis. These components are related to syntactic violations. Besides, this work also proposed an approach calculating the cross-correlation function between ERSP and stimuli. Though this method has several weaknesses, it still proves strength in analyzing brain working mechanism.

REFERENCES


Aparicio, X., Midgley, K. J., Holcomb, P. J., Pu, H., Lavaur, J. M., & Grainger, J. (2012). Language Effects in Trilinguals: An ERP Study. *Front Psychol, 3*, 402. doi:10.3389/fpsyg.2012.00402

Bastiaansen, M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *J Cogn Neurosci, 22*(7), 1333-1347. doi:10.1162/jocn.2009.21283

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci, 25*, 151-188. doi:10.1146/annurev.neuro.25.112701.142946

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage, 26*(1), 221-233. doi:10.1016/j.neuroimage.2005.01.032

Correia, J. M., Jansma, B., Hausfeld, L., Kikkert, S., & Bonte, M. (2015). EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations. *Front Psychol, 6*, 71. doi:10.3389/fpsyg.2015.00071

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Res, 1158*, 81-92. doi:10.1016/j.brainres.2007.04.082

Dikker, S., Rabagliati, H., & Pylkkanen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*(3), 293-321. doi:10.1016/j.cognition.2008.09.008

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn Sci, 6*(2), 78-84. doi:10.1016/s1364-6613(00)01839-8

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol Rev, 91*(4), 1357-1392. doi:10.1152/physrev.00006.2011

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language, 43*(3), 476-507. doi:10.1006/jmla.2000.2709

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *J Psycholinguist Res, 31*(1), 45-63. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11924839

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Brain Res Cogn Brain Res, 1*(3), 183-192. doi:10.1016/0926-6410(93)90026-2

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research, 1146*, 50-58. doi:10.1016/j.brainres.2006.08.038

Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *Behav Brain Sci, 23*(1), 1-21; discussion 21-71. doi:10.1017/s0140525x00002399

Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Curr Opin Neurobiol, 16*(2), 240-246. doi:10.1016/j.conb.2006.03.007

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*(4), 556-568. doi:Doi 10.1162/089892900562336

Hahne, A., & Jescheniak, J. D. (2001). What's left if the Jabberwock gets the semantics? An ERP investigation into semantic and syntactic processes during auditory comprehension. *Cognitive Brain Research, 11*(2), 199-212. doi:Doi 10.1016/S0926-6410(00)00071-9

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci, 8*(5), 393-402. doi:10.1038/nrn2113

Hinojosa, J., Martín-Loeches, M., Casado, P., Muñoz, F., & Rubia, F. (2003). Similarities and differences between phrase structure and morphosyntactic violations in Spanish: An event-related potentials study. *Language and Cognitive Processes, 18*(2), 113-142. doi:10.1080/01690960143000489

Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology, 74*(3), 337-346. doi:10.1016/j.biopsycho.2006.09.003

Kielar, A., Meltzer, J. A., Moreno, S., Alain, C., & Bialystok, E. (2014). Oscillatory responses to semantic and syntactic violations. *J Cogn Neurosci, 26*(12), 2840-2862. doi:10.1162/jocn_a_00670

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage, 105*, 507-524. doi:10.1016/j.neuroimage.2014.11.016

Klimesch, W., Doppelmayr, M., & Hanslmayr, S. (2006). Upper alpha ERD and absolute power: their meaning for memory performance. In *Event-Related Dynamics of Brain Oscillations* (pp. 151-165).

Knosche, T. R., Maess, B., & Friederici, A. D. (1999). Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topography, 12*(2), 75-87. doi:Doi 10.1023/A:1023442426799

Krause, C. M. (1999). Event-related Desynchronization (ERD) and Synchronization (ERS) During Auditory Information Processing. *Journal of New Music Research, 28*(3), 257-265. doi:10.1076/jnmr.28.3.257.3113

Kubota, M., Ferrari, P., & Roberts, T. P. L. (2003). Magnetoencephalography detection of early syntactic processing in humans: comparison between L1 speakers and L2 learners of English. *Neuroscience Letters, 353*(2), 107-110. doi:10.1016/j.neulet.2003.09.019

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. doi:10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Mem Cognit, 11*(5), 539-550. doi:10.3758/bf03196991

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically Based Sentence Processing Classes - Evidence from Event-Related Brain Potentials. *Journal of Cognitive Neuroscience, 3*(2), 151-165. doi:DOI 10.1162/jocn.1991.3.2.151

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci, 12*(6), 718-724. doi:10.1038/nn.2331

Soskey, L., Holcomb, P. J., & Midgley, K. J. (2016). Language effects in second-language learners: A longitudinal electrophysiological study of spanish classroom learning. *Brain Res, 1646*, 44-52. doi:10.1016/j.brainres.2016.05.028

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., . . . Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage, 30*(4), 1414-1432. doi:10.1016/j.neuroimage.2005.11.002

Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J. M., Magyari, L., . . . Bastiaansen, M. (2012). Beta oscillations relate to the N400m during language comprehension. *Hum Brain Mapp, 33*(12), 2898-2912. doi:10.1002/hbm.21410

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *J Acoust Soc Am, 75*(6), 1866-1878. doi:10.1121/1.390988

Willems, R. M., Oostenveld, R., & Hagoort, P. (2008). Early decreases in alpha and gamma band power distinguish linguistic from visual information during spoken sentence comprehension. *Brain Res, 1219*, 78-90. doi:10.1016/j.brainres.2008.04.065