Master of Science

Dominancy in Speech Comprehension based on Phoneme-related Representation of Electroencephalography

The Graduate School

of the University of Ulsan

Department of Biomedical Engineering

Le Thi Trang

Dominancy in Speech Comprehension based on Phoneme-related Representation of Electroencephalography

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

Le Thi Trang

Department of Biomedical Engineering

University of Ulsan, Korea

June 2020

Dominancy in Speech Comprehension based on Phoneme-related Representation of Electroencephalography

This certifies that the master's thesis

of Le Thi Trang is approved.

Committee Chair, Dr. Young-joon Chee

Committee Member, Dr. Jihwan Woo

Committee Member, Dr. Kyo-in Koo

Department of Biomedical Engineering

University of Ulsan, Korea

June 2020

ACKNOWLEDGMENT

First of all, I would like to express my special thanks to Professor Woo Jihwan, who offered me the opportunity to be a master's student in the CNE lab, which I have had a great time during these two years. I always appreciate his gentility and carefulness for not only me but also for all his students.

Secondly, I also would like to give my thanks to my labmates as well as my Vietnamese friends both in Ulsan and Vietnam who have always been kind to me and help me a lot in my beginning days in Korea. I thank all the trips we went to, all the chat we have had, all the birthday events, the dinners and regular occasions we spent together. I believe I have had beautiful memories that I will always remember and thank for. Without them, I could not be adaptive to the new life in another country as well as now.

Next, my acknowledgement to my family is perpetual. I gratefully acknowledge the support from my family, especially my mom and my sister for their endless encouragement, base on which I can pursuit academic study without worry.

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

Thank you,

Le Thi Trang

Contents

ABSTRACT

Verbal communication comprises retrieval of both semantic and syntactic information elicited by various kinds of words in a sentence. Content words, such as nouns and verbs, convey essential information about the overall meaning (semantics) of a sentence whereas function words, such as prepositions and pronouns, carry less meaning and support the syntax of the sentence. The aim of this study is to reveal neural correlates of differential information retrieval processes for both content and function words by examining electroencephalography during spoken-sentence comprehension. The distinction between spatiotemporal patterns of cortical responses to different parts of speech provides neurological evidence for the variable effects on comprehension of omitting certain types of words within a sentence. The findings of this study may provide insight to the different contributions of certain types of words over others in overall sentence understanding based.

List of figures

List of tables

1. INTRODUCTION

1.1. Research basis and objective

Sentence comprehension is crucial for effective communications among humans. In most languages, words are grammatically categorized into several parts of speech: noun, verb, pronoun, preposition, adverb, conjunction, particle, and article. Different parts of speech are thought to induce differential information retrieval processes [1], as they have varying degrees of importance in sentence comprehension. The parts of speech may be divided into two major groups: content words and function words. The content words, such as nouns, verbs, adjectives, and adverbs, play an important role in sentence comprehension by conveying irreplaceable critical meaning. In contrast, function words, such as prepositions and pronouns, provide continuity in the sentence without holding much meaning on their own [2]. Consequently, we hypothesize that content words are more important than function words for sentence comprehension. Within content words, nouns and verb may convey different amount of information. In addition, the thematic role assignment that determines whether a verb's argument is an object or subject of the action provides another critical aspect of sentence comprehension [3, 4]. Accordingly, within nouns, object words and subject words may play distinct roles in sentence comprehension. While listening to spoken sentences, listeners' attention applied to each component may vary as a function of its importance for understanding the meaning of the whole sentence. This study aimed to provide neurophysiological correlates of such dominance differences in linguistic components for sentence comprehension.

Speech processing entails numerous steps that happen within short time intervals, around a hundred milliseconds [5, 6] as the speech signal unfolds over time. Since electroencephalography (EEG) can measure brain activity with excellent temporal resolution, it has been widely used to investigate language and speech processing [6, 7]. Specifically, event-related potentials (ERPs) can reveal variable cognitive processes in a time- and phase-locked temporal pattern [5]. Although traditional ERP techniques rely on repeated measures of isolated sensory events (e.g., isolated syllables: [8-10] recent advances in EEG analysis made it possible to parse ERP components elicited by individual segments (e.g., phonemes and word onsets in a sentence) from continuous speech stimuli (e.g., sentences). Such analyses are based on computing cross-correlation between continuous EEG signals and time course of speech stimuli (e.g., temporal envelopes) [11-13]. This study utilizes a cross-correlation-based method that is developed to parse early cortical ERP components (e.g., N1) following linguistic events in naturally spoken continuous speech.

Several methods for extracting acoustic events from continuous speech have been suggested. The most common way was using temporal envelopes [14, 15] assuming that the envelope provides important acoustic information for sentence understanding [16]. However, [17] contends that phoneme-level processing was encoded in low-frequency cortical oscillations and the phonemic model outperformed the envelope model in predicting neural responses to speech. In addition, phoneme information is posited to reveal brain activation under spoken sentence perception task [18-20]. Consequently, the current study utilizes phoneme-onset impulse trains as the event time course of continuous speech for the cross-correlation based EEG analyses.

Previous neuroimaging studies investigated cortical networks underlying several aspects of sentence comprehension. A temporo–frontal network has been argued as a major region for sentence comprehension within which temporal regions identify linguistic events whereas frontal regions build syntactic and semantic relations [21, 22]. Although scalp EEG does not provide an optimal method to investigate the cortical sources of activities, a few studies examine syntactic and semantic processing using EEG by utilizing syntactic and/or semantic violations within sentences [23-26]. Instead of manipulating certain components of natural speech stimuli, we selectively omit linguistic components from indexed impulse trains of phoneme onsets in a systematic way. Also, instead of estimating cortical sources of activities, we seek well-known ERP components (e.g., N1) from the combination of morphology and topography analyses. Because we hypothesize that differential dominance of linguistic components for sentence comprehension will be reflected in the strength of attention applied to such components, we mainly focus on auditory N1 component which is strongly modulated by attention [27, 28].

1.2. Basic brain anatomy and function

The cerebral cortex of the human brain is traditionally divided into four lobes: frontal, temporal, parietal and occipital purely based on anatomical aspect. Although, the most brain functions are executed by cooperation of many different regions across the brain, it has been known that each lobe carries out a certain number of that. Figure 1 presents four lobes, and the functions which are claimed to be supported by the corresponding lobes, motor cortex, and sensory cortex are also parts of the frontal lobe.

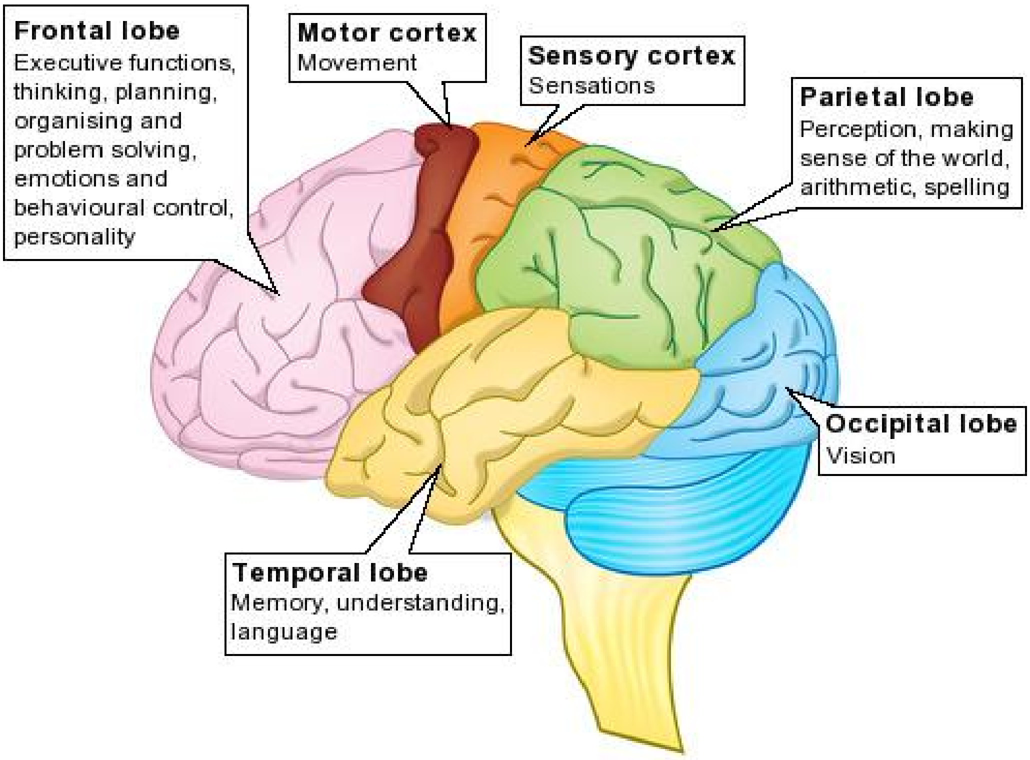

Figure 1: Brain's lobes and their function[1]

[1] www.headway.org.uk

Brodmann's map is a cytoarchitectonic map of the human brain which was first introduced in 1909 by Korbinian Brodmann, a German neurologist. Brodmann parcellated the cerebral cortex into "areas" due to the differences in anatomical aspects among "areas". Brodmann indicated his numbering system homologies in comparative brain neuroanatomy of different mammals. In according to this, he divided the human brain into 52 areas, which were numbered from 1 to 52. In terms of the human brain, only 43 cortical areas belonging to 11 regions are displayed in the Brodmann's map while the number of that for monkeys and great apes are 30. The reason for that, as Brodmann explained, is the absence of some areas in the human brain while they are well observed in other mammalian species[29]. The introduction of novel instruments and techniques for neuroimaging facilitates linking architectonic units and their function that make Brodmann's map more popular. In fact, there resist limitations to some extent, one of that is the intersubject variability of cytoarchitectonic areas. Even though there are doubts about the Brodmann's map, the map has still been commonly used in a large number of researches of neuroscience filed until now. Figure 2 shows the Brodmann's map of the left hemisphere of the human brain.

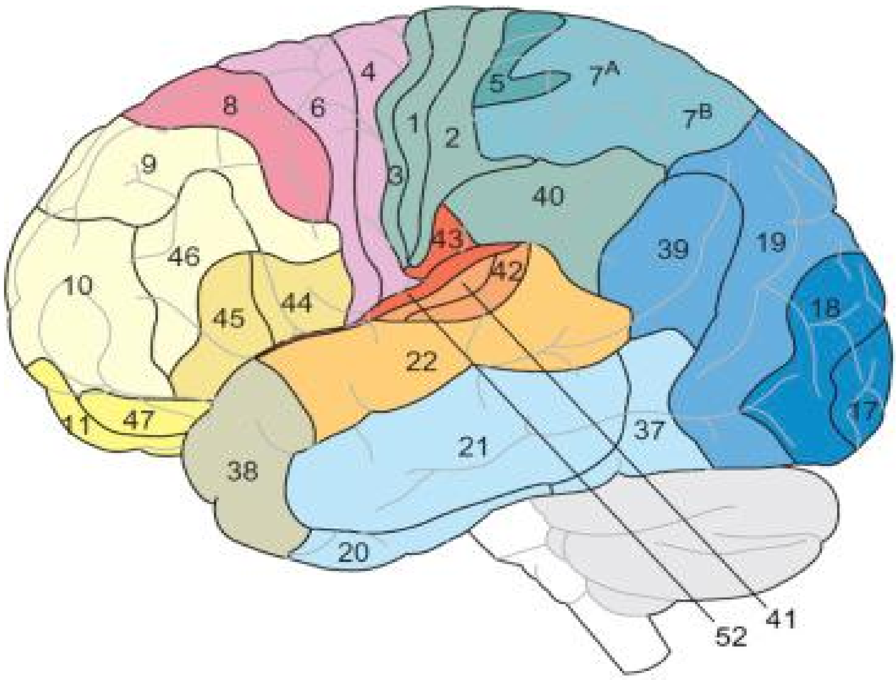

Figure 2: The Brodmann's map illustration for the left hemisphere[2]

[2] 30. Gage, N.M. and B.J. Baars, *Decisions, Goals, and Actions*, in *Fundamentals of Cognitive Neuroscience*. 2018. p. 279-319.

1.3. The language-relevant regions

As shown in Brodmann's map, the human brain was parcellated into different areas, which have difference in structure aspects. However, it is also known that each area supports wide range of functions, such as: language processing, working memory, visual processing, etc. Moreover, there exist connections between areas which contribute to that function. With respect to language processing, the language-relevant areas consist of Broca's area. Wernicke's area, the middle temporal gyrus (MTG) and the inferior parietal and angular in the parietal lobe; those areas and connections between them are coded in difference color legend in the Figure 3. The frontal gyrus, where Broca's area is located, is argued to support variety of language processing aspects: phonology and articulation, syntactic and semantic processing, etc. Additionally, the temporal cortex, which covers the primary, secondary auditory cortex (BA 41/42), Wernicke's area (BA 22), is known to be in sound and speech processing as well as the processing of words and phrases. In addition to the left frontal and temporal cortex, the parietal cortex is also claimed to be related to language processing: phonological working memory and storage [31]. In terms of connections between language-related regions, dorsal and ventral pathways are known to connect the frontal cortex and the temporal cortex. Regrading to this concept, the dorsal and ventral pathways have been suggested to support the auditory-motor integration and sound-to-meaning mapping, respectively. The direction symbols are argued to be in the feed-forward and back projections concept. However, the structure connections were defined only based on the function of particular regions which they connect [32]. Although, the findings have been supported by several researches and become common in the domain of human language processing. Figure 3 presents the language-relevant regions and the language-related pathways between the inferior and temporal cortex.

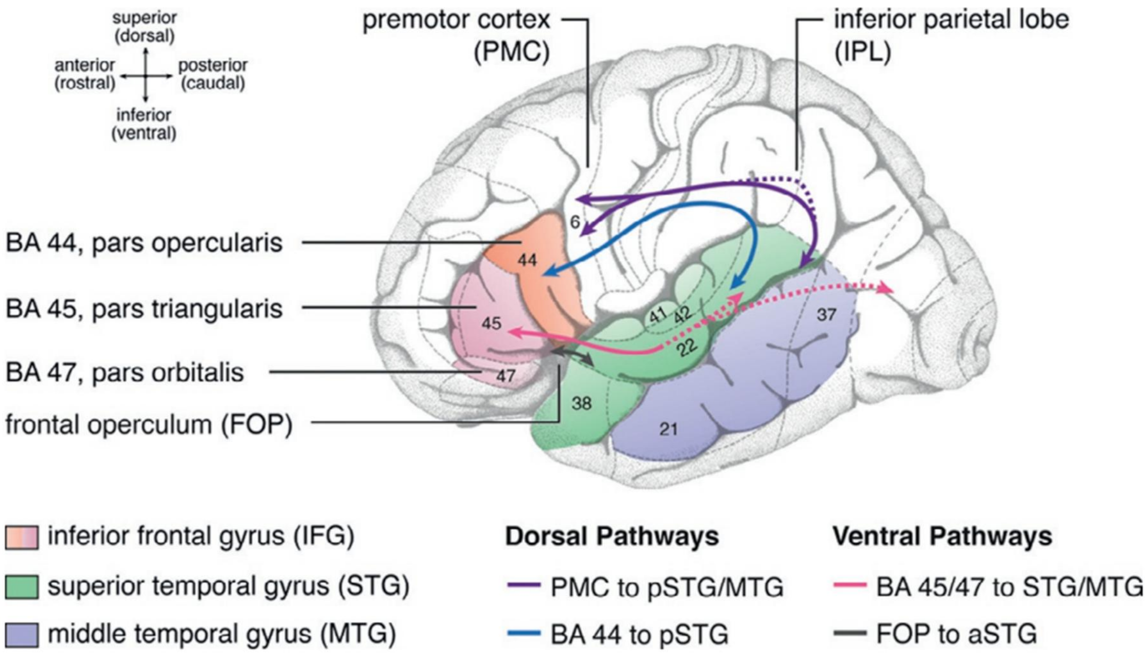

Figure 3: The language-relevant brain regions and their connection[3]

[3] 31. Friederici, A.D., *White-matter pathways for speech and language processing.* Handb Clin Neurol, 2015. **129**: p. 177-86.

1.3.1.Broca's Area

Broca's area is the brain region which is located in the frontal lobe of the dominant hemisphere, usually in left side of the human brain, which is named after Paul Broca, an anatomist, anthropologist and surgeon. The area was defined in terms of the pars opercularis and pars triangularis as BA (Brodmann's Area) 44 and 45 areas corresponding to Brodmann's map. This area was first discovered in 1861 since Broca presented the behavioral and neuropathological data of a man patient, who lost their language speaking ability but could understand language because of severe damage or primary lesion in some sub-regions in left frontal lobe, which became knowns as Broca's area. In concept of that, Broca's areas were considered to involve in the production of speech as Broca claimed [33, 34]. In several years, many researches also indicated the correlation between the impairment in the left inferior frontal gyrus, where Broca's area is located, and the disturbance in speech comprehension. For that reason, Broca's area is now considered to be relative to both speech production as well as speech comprehension. A homologous area in the right hemisphere is claimed to entail the processing of prosody of speech. Figure 4 shows the Broca's area in concept of Brodmann's map with different color code legends for BA 44 and BA 45.

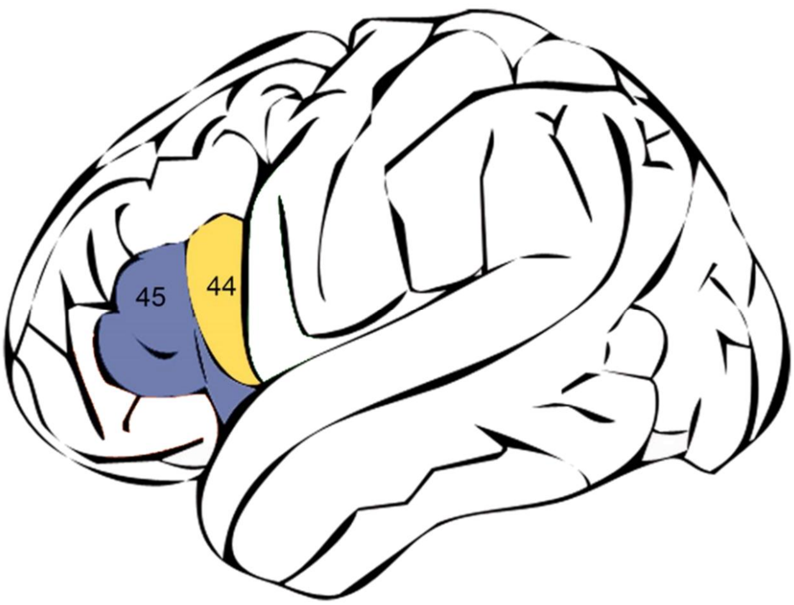

Figure 4: Broca's area in terms of Brodmann's map[4]

[4] 35. https://en.wikipedia.org/wiki/Wernicke%27s_area. *Broca's area*.

1.3.2.Wernicke's Area

Wernicke's area is a region of the brain which is located in the posterior section of the superior temporal gyrus (STG) of the dominant hemisphere of the brain, which is usually in the left. This area is in the vicinity of the auditory cortex, which is a part of auditory system and has charge of auditory information processing in humans. According to Brodmann's map, Wernicke's area is thought to reside in BA 22 which is in the left hemisphere of 95% of right-handed population and 60% for the left-handed populations [35]. Wernicke's area was first described by a neurologist named Carl Wernicke in 1874 when he identified the similar type of problem with that of Broca's experiment in his patients, who were able to speak but has language comprehension disorder. In cooperation with Broca's area, Wernicke's area is a part of the cerebral cortex that is associated with the spoken and written language understanding. Damage caused to Wernicke's area results in language aphasia which is known as "Wernicke's aphasia". The homologous area, which is equivalent to the Wernicke's area in the non-dominant hemisphere, is considered to play a role in ambiguous words processing. Figure 5 below illustrates the location of Wernicke's area and its location correlation with Broca's area.

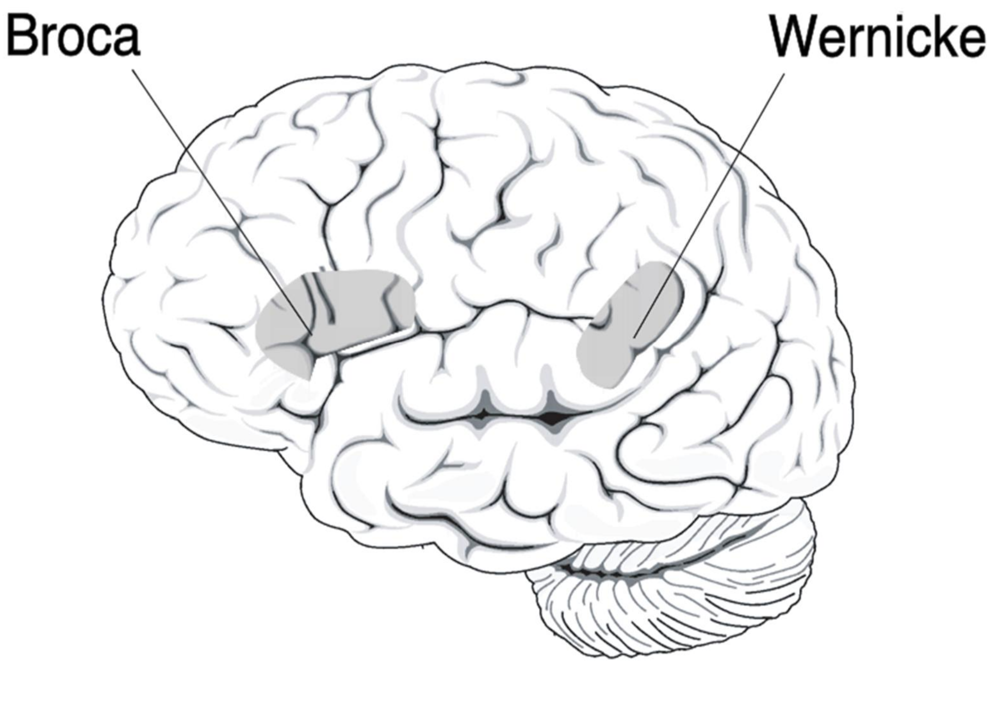

Figure 5: Location of Wernicke's areas in relation with that of Broca's area[5]

[5] 35. Ibid.

1.4. Electroencephalography

Electroencephalography (EEG) is a technique for recording the electrical activity of the brain. It was first recorded on human in 1924 by Hans Berger, a German physiologist and psychiatrist. EEG, a non-invasive method, measures voltage fluctuations within neurons of the brain by using electrodes cap placed on the scalp as shown in Figure 6 below. Usually, an EEG signal can be decomposed into component of different frequencies. For example, Figure 7 illustrates the typical brain wave of sleep and wakefulness of a brain person recorded by using EEG.

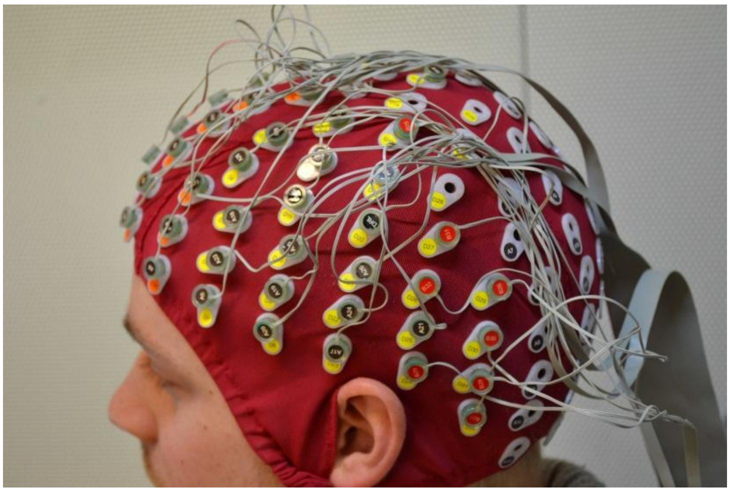

Figure 6: An EEG recording setup[6]

Typically, it is known that most cognitive processes of human occur within ten to hundreds of milliseconds which make EEG is an exceptional tool for studying that of processes. Additionally, EEG directly measures neural activity which is direct reflection of biophysical phenomena occurred in human brain. Moreover, EEG is a multidimensional signal which facilitate for stating and testing the hypotheses both in neurophysiology and in psychology. In contrast, EEG remains limitation as its poor spatial resolution compared to other neuroimaging techniques as fMRI, CT, PET. In addition, EEG also has low level of signal-to-noise ratio (SNR). To improve the SNR, ERPs (event-related potentials) which is obtained by averaging out the EEG are usually used instead of the raw EEG signal. Besides that, ERPs not only hold the quality of EEG but also gain another advantage over EEG as time-locked to stimulus, which is an essential characteristic of a specific cognitive event [36]. Base on that feature, ERPs include 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, also known as N1. These components are often variable but particular for the

[6] https://www.psychologie.uzh.ch/de/bereiche/ehem/plasti/Labor.html

cognitive processing of the stimulus [37]. Therefore, ERPs have been widely used in neuroscience, cognitive psychology, and cognitive science, etc. Figure 8 displays an example of ERP and its components measured in a neurolinguistics research.

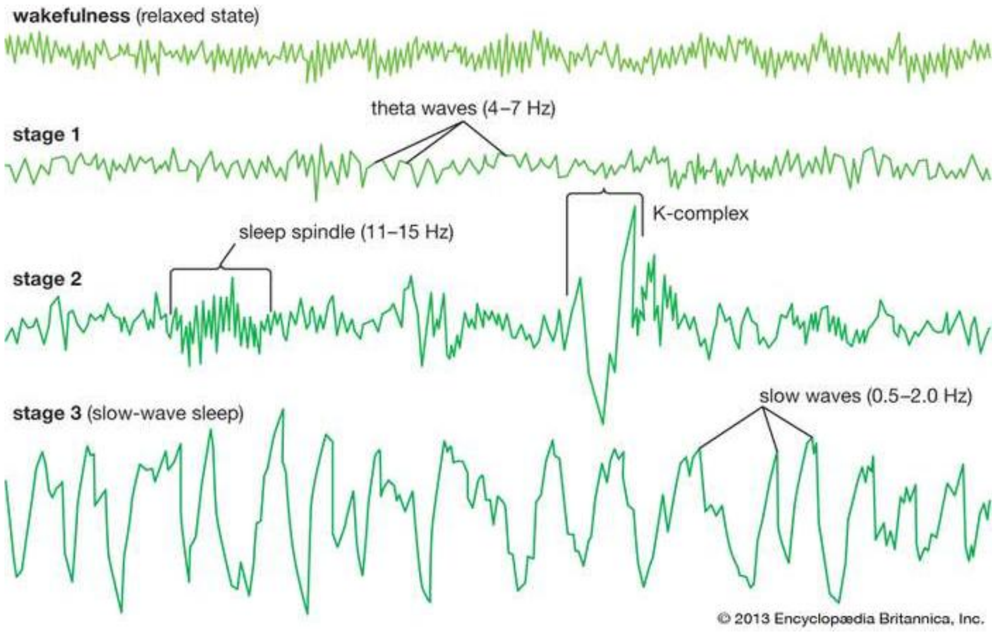

Figure 7: An example of typical brain wave of an EEG of sleep and wakefulness[7]

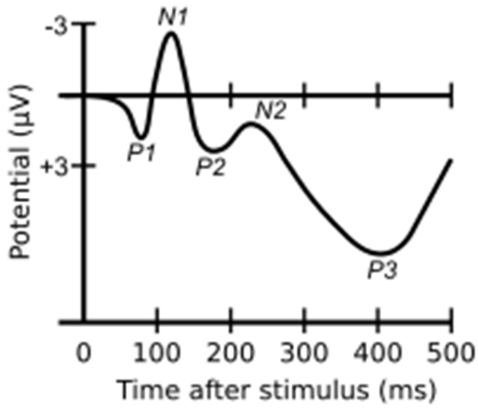

Figure 8: An example of ERPs components[8]

[7] https://www.britannica.com/science/electroencephalography

[8] 37. https://en.wikipedia.org/wiki/Event-related_potential. *Event-related potential*.

2. BASIS THEORY

2.1. EEG recording system

2.1.1. Experiment diagram

Thirteen normal-hearing adult native speakers of English were involved in the experiment. Ten sentences spoken by an American male were used as auditory stimuli. The experiment was conducted in a soundproof booth, where the participants seated facing a computer monitor and four speakers placed at the front of them at a distance of 1 m. The experiment setup details are shown in Figure 9 as below.

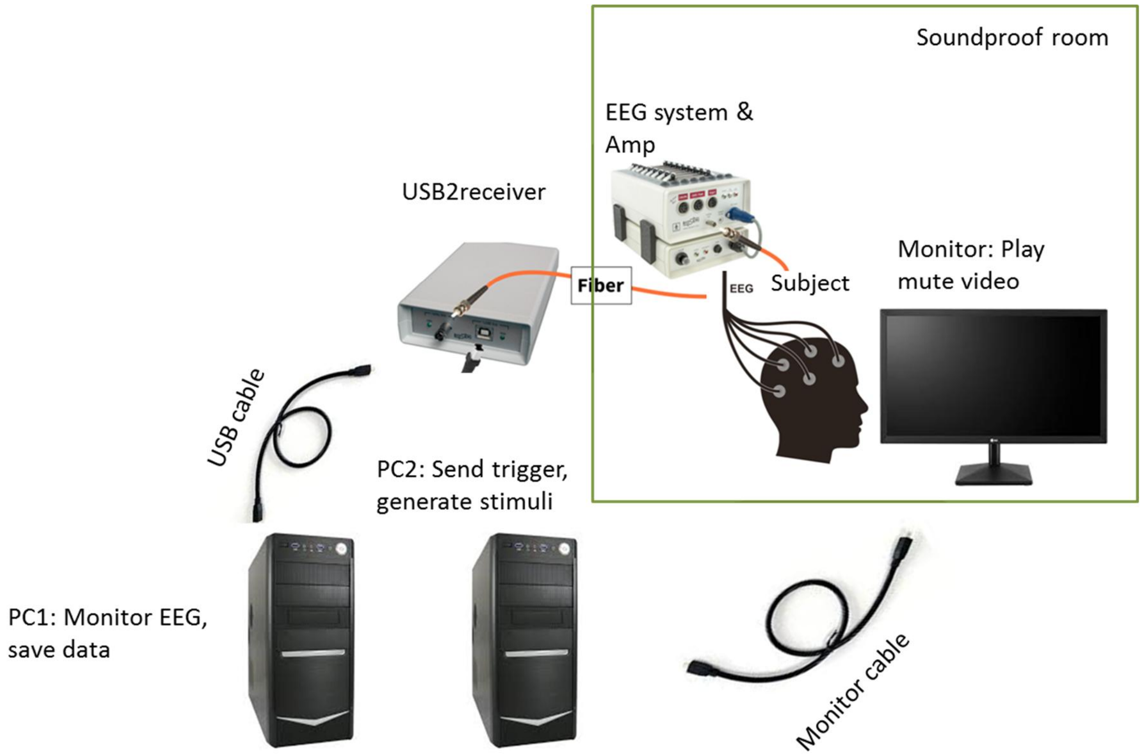

Figure 9: EEG recording diagram

2.1.2.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 Psychtoolbox and Matlab environment. The routines at the core of Psychtoolbox allow synchronization, support sub-millisecond timing and capture low-latency audio which are important for cognitive process such as language processing, which is usually occur at few milliseconds after stimuli onset.

2.1.3.Experiment procedure

In preparation for the experiment, participants were asked to wear 64 Ag-AgCl electrode holders cap which holders are distributed as 10/20 layout; then all electrodes were placed in corresponding holders. After that, electrode holders were filled with electrode gel with a syringe; electrodes impedance was monitored via ActiView software installed in PC1 (Figure 9) to ensure that the impedance was kept in the range from -50 μV to 50 μV.

Ten English sentences were used as auditory stimuli; each sentence has length not excess 2.5 seconds. Speech sentences were played by a speaker placed right in front of the participants. During the experiment, all sentences were played in a random order with an interval of 3 second. Each sentence was played 100 times, which resulted in 1000 trials in total. Participants were asked to follow a silent movie presented on the monitor to ensure passively listening to the speech sound. In the course of the experiment, participants were also required to restrict movements in order to minimize motion artifacts. EEG signals were monitored and saved in real time using ActiView display software run on PC1. All the movements of subjects and corresponding trial were noted down by manually monitored though camera in the soundproof room.

2.2. Stimuli analysis

2.2.1.Sentence structure

In linguistic, English words are grammatically categorized into parts-of-speech: noun, verb, pronoun, preposition, adverb, conjunction, particle and article. The information retrieval process can be affected by parts-of-speech. Thus, these components were thought to have varying degrees of importance in sentence perception. Part-of-speech (POS) may be divided into two major groups: content words and function words. The content words, which consist of nouns, verbs, adjectives, and adverbs play the important role in sentence comprehension while function words such as preposition, pronoun, etc. have a little or no meaningful effect. In other words, content words provide the most important information while function words link these words together. Regarding to this concept, noun and verb provide the information about person, place or thing, and action or state of action, respectively. Adjective usually describes the noun while adverbs tell us how, when or where the noun happens. In contrast, the function words' components just play role as grammatical words such as: referring to other nouns using the pronouns, connecting words using conjunctions. It is claimed that, content words are usually stressed in conversation while function words got no stress. In addition to that, thematic role which determines whether a verb's argument is an object or subject of the action provides another critical aspect of sentence comprehension [4]. In conclusion, early detecting parts-of-speech facilitate spoken sentence processing. Table 1 and Table 2 below show the traditional parts of speech in English language and a part of alphabetical list of POS tag [38], respectively.

Table 1: Traditional parts of speech

Word categories	Meaning	Example
Noun	Name of things	Boy, cat, truth
Verb	Action or state	Become, hit
Pronoun	Used for noun	I, you, we
Adverb	Modifies Verb, Adjective, Adverb	Sadly, very
Adjective	Modifies noun	Happy, clever
Conjunction	Join things	And, but, while
Preposition	Relation of N	To, from, into
Interjection	An outcry	Ouch, oh,

Table 2: A part of alphabetical list of POS tag used in Penn Treebank

Tag	Description
VB	Verb, base form
VBD	Verb, past tense
NN	Noun, singular or mass
NNS	Noun, plural
JJ	Adjective
IN	Preposition or subordinating conjunction
DT	Determiner

2.2.2. Sentence structure parser

A natural language parser is a program that works out the grammatical structure of sentences. In other words, a language parser can analyze a sentence into part-of-speech, for instance, which word are noun or verb and which word is subject or object of a verb. Nowadays, there are many projects about sentence structure parser have been developed. One of that is Natural Language Toolkit, which is a leading platform in Python which provide an API for completing the task of labeling POS for components in English sentence [39]. Another project is also used for sentence parser is the Stanford Parser, a Java package, which can be adapted to many languages such as: English, Chinese, German, etc. [40]. The parser uses the Penn Treebank, which produced approximately 7 million words of POS tagged text. Figure 10 below illustrates the POS tags of a sentence, "Maple syrup is made from sap", analyzed using the Stanford Parser, in which, each component was followed by their POS tag.

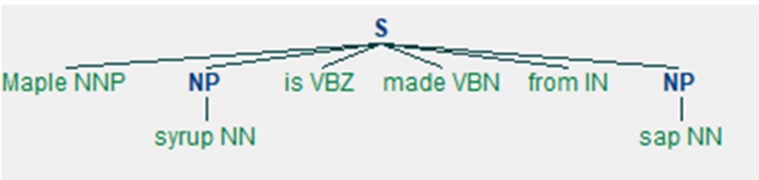

Figure 10: An example of POS tag of a sentence

2.2.3.Phoneme impulse train

Beside the grammatical aspect, speech sentence also includes the phoneme information. A phoneme is a unit of sound that distinguishes one word from another in a particular language. Usually, one or more phonemes compose one word. One can separate each phoneme in an uttered English sentence by its onset time using the Praat software, which is used for speech analysis in phonetics [41]. Figure 11 displays an example of doing phonetics using Praat, in which each phoneme onset time is presented by a blue vertical line, and the corresponding time point is shown right above that.

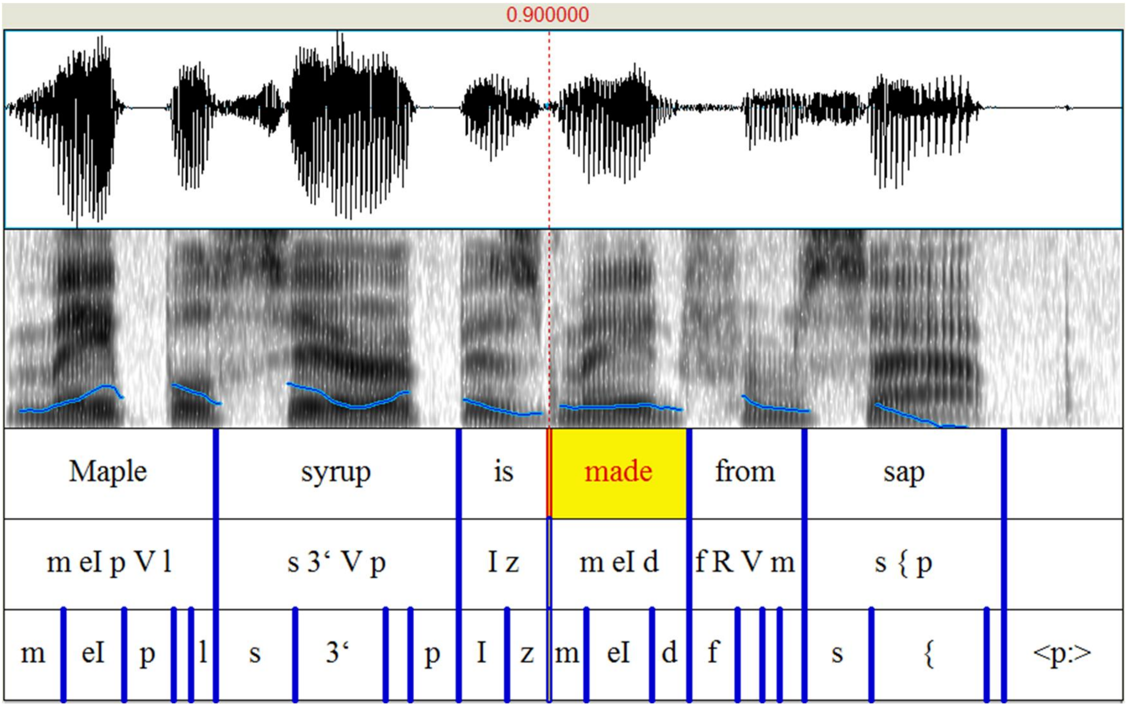

Figure 11: An example of phoneme analysis using Praat

Based on that, phoneme impulse train was obtained for each stimuli sentence. To examine in what degree each POS component involves in sentence comprehension, the phoneme relating to that component was omitted to make the component-excluded phoneme impulse train. The components were examined includes content words, function words, nouns, verbs, objects and subjects, which resulted in 6 cases of component-excluded: content words-, function words-, nouns-, verb-, objects-, and subjects-excluded case., respectively. Each component was compared to the whole-sentence with the other in pair, which are content word vs. function words, noun vs. verb, and object vs. subject. The reason for that is the reciprocal relationship between the two components in each pair to proper grammar of sentence. Importantly, there were no significant differences (Wilcoxon signed rank test, $p>.05$) in the number of phonemes in the case of each pair (i.e. noun vs. verb; object vs. subject).

2.3. Data processing procedure

2.3.1. EEG recording and processing

EEG signals were recorded using a 64-channel EEG system (BioSemi Co., Netherlands) at 2048 Hz sampling rates. Then the pre-processing procedure was conducted using EEGLAB, which is an interactive Matlab toolbox for processing continuous and event-related EEG, MEG. Firstly, the raw EEG signals were pre-processed using an average reference to remove 40dB noise in the data. Secondly, each channel data was filtered using 0.5-57 Hz 5th order Butterworth filter. Usually, EEG signals are usually mixed recorded with other biosignals such as EMG, ECG, eye-blinking, etc. as shown in Figure 12.

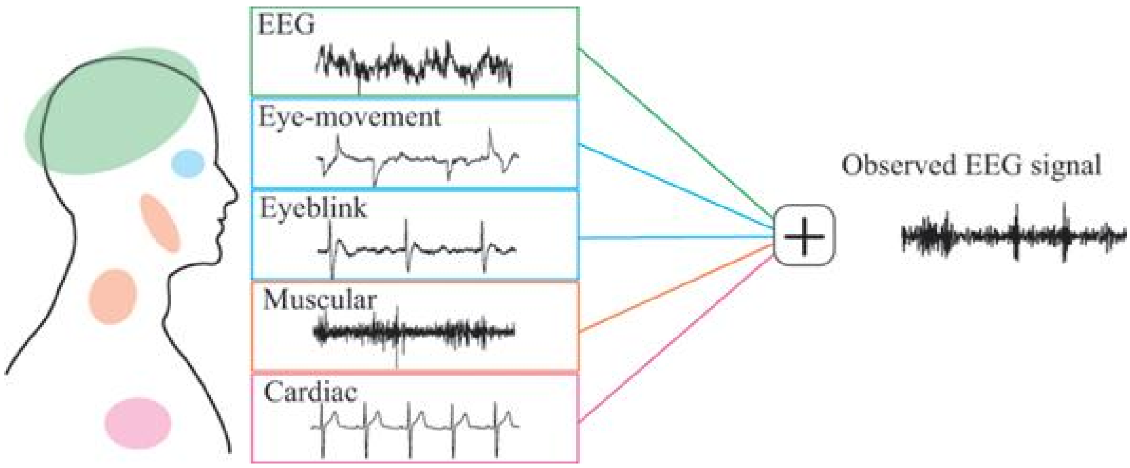

Figure 12: Example of components mixed in EEG signals

To unmix the noise components in EEG, the Independent Component Analysis (ICA), widely-used blind source separation technique, was implemented as the third step in the pre-processing procedure. With ICA, EEG signals are separated into a number of components. Figure 13 below shows an example of a component, electrooculogram (EOG), which was separated from EEG using ICA. Typically, noise components were rejected through visual inspection with the topographic map (Figure 13).

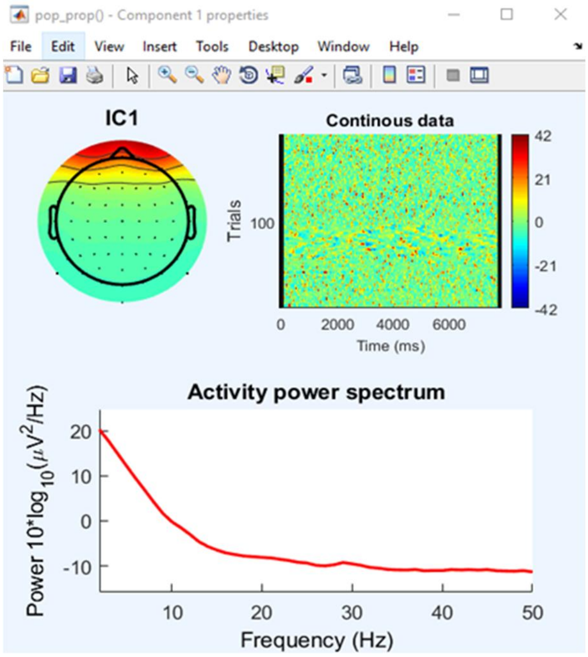

Figure 13: Example of a component analyzed using ICA (EOG)

According to that, EOG component was rejected. The EEG signals were then analyzed using epoch length of 3 s from 0.5 s prior to stimulus onset to 2.5 s after stimulus onset. Epochs that exceeded 50 µV were excluded. Then, the epochs were concatenated and filtered using 1-15 Hz 5th order Butterworth bandpass filter.

In summarize, EEG signals recorded during the experiments were preprocessed following below steps:

- Re-reference
- Filter (0.5-57, order: 5th)
- Independent components analysis
- Remove Eye artifact based on topography
- Epoch based on trigger (-0.5-2.5 s)
- Remove out of bounds epochs (-50-50 µV)
- Filter concatenated epochs (1-15 Hz, order 5th)

2.3.2.EEG analysis

The ERP were computed by averaging all epochs from each sentence at each EEG electrode. The ERP baseline was corrected by subtracting the average amplitude between -200 and 0 ms relative to stimulus onset. An example of the ERP obtained from the sentence “Maple syrup is made from sap” is shown in Figure 14C. Figure 14A and B show an illustration of phoneme-onset impulse trains of the sentence “Maple syrup is made from sap” for a whole-sentence and for a noun-exclusion case, respectively. There were no significant differences (Wilcoxon signed rank test, p>.05) in the number of phonemes in the case of each pair (i.e. noun vs. verb; object vs. subject).

Cross-correlation coefficients were computed between phoneme-onset impulse trains and their corresponding average EEG signals as a function of lag that ranged from -200 to 700ms (where positive sign indicates that the EEG signal lags the phoneme onset train). Then the cross-correlation function subtracted mean value of cross-correlation coefficients and passed through a 1-Hz high-pass filter. Figure 14 D illustrates an example cross-correlation function obtained using the phoneme impulse train shown in Figure 14 A and its corresponding EEG signals show in in Figure 14 C.

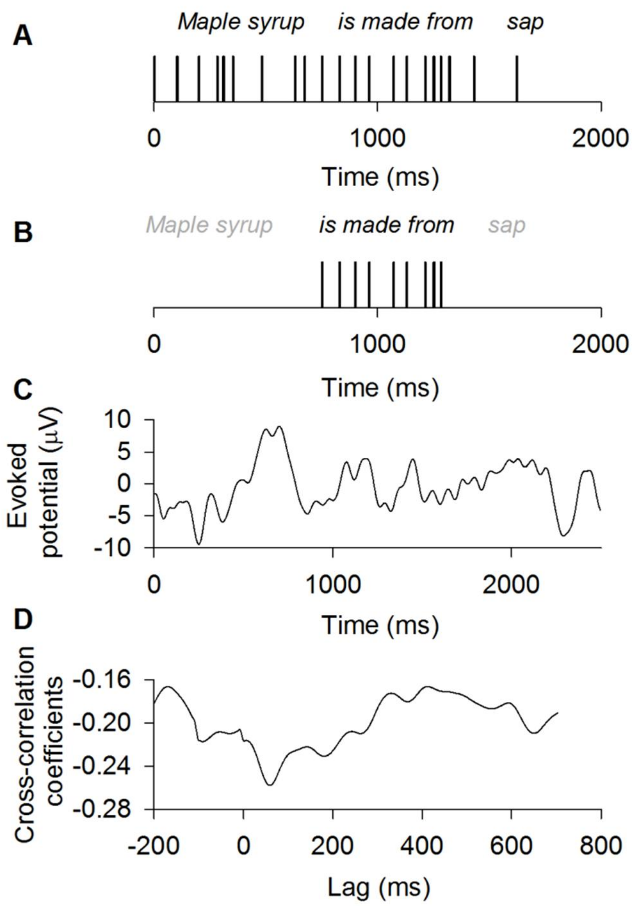

Figure 14: An example of (A) phoneme impulse train for whole-sentence, (B) noun-excluded case, (C) ERP and (C) cross-correlation between (A) and (C)

3. RESULTS

Figure 15 shows the EEG vs. phoneme onset-train cross-correlation functions averaged across all subjects and all the 10 sentences at left fronto-central electrode sites of AF7, F5, and F7 for function- vs. content word-exclusion cases (A), verb- vs. noun-exclusion cases (B), and subject- vs. object-exclusion cases (C). In each panel, the exclusion cases are represented by black or red lines whereas the dark gray line shows the whole-sentence case. For the whole-sentence case, the traditional P1-N1-P2 ERP morphology was clearly observed with peaks at 152.3, 199.2, and 285.2 ms, respectively. The cross-correlation coefficients at such peaks (e.g., N1) were significantly reduced in several exclusion cases. For instance, the N1 component in content word-exclusion case has been significantly reduced compared to the whole-sentence case in brain regions related to language processing (Table 3, Wilcoxon signed-rank test, $p < 0.05$) whereas function word-exclusion case did not reveal such a significant difference (See Figure 15 A, the panel F7).

Table 3: Channels showed significant difference ($p<0.05$) in N1 peak amplitude where N1 peak amplitude was reduced in each component-exclusion case compared to whole-sentence case

	Content word-exclusion	Function word-exclusion	Noun-exclusion	Verb-exclusion	Object-exclusion	Subject-exclusion
Number of electrodes	4	3	10	0	31	4
Electrodes	C3, C5, Cz, T7	P9, TP8, PO8	Fpz, AFz, Fz C3, C5, CPz, Cz, Fp2, C2, P2		AF3, F1, F3, F5, FC3, FC1, AFz, Fz, FCz C1, C3, CP3, CP1, P9, PO3, CPz, Cz, O1, Oz, F2, F4, F6, FC6, FC4, FC2, C2, C4, C6, CP4, CP2, P8	T7, Fpz, AFz, P2

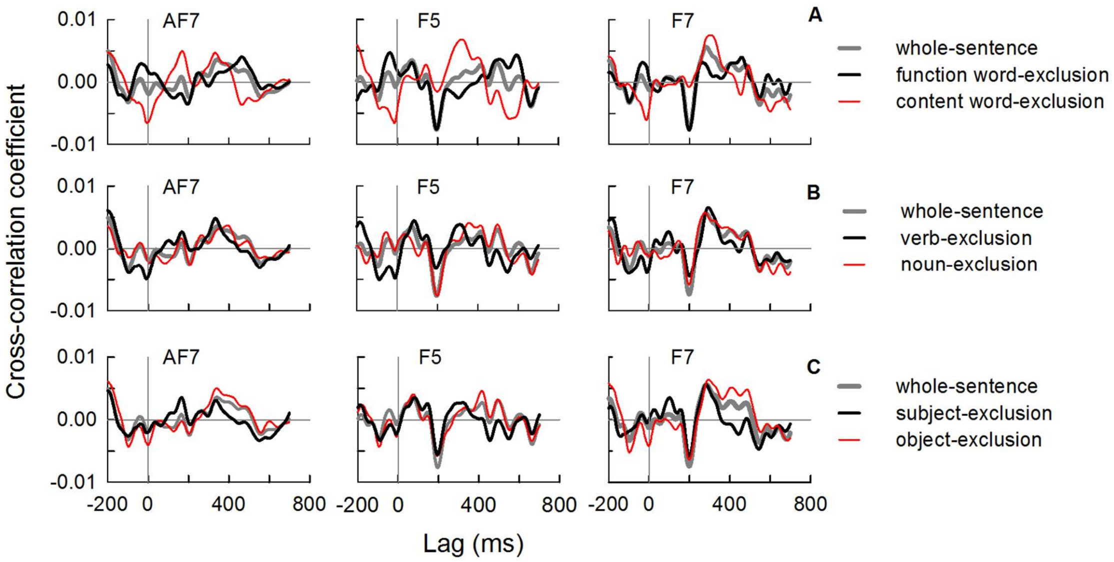

Figure 15: Example of cross-correlation functions at the typical electrodes of AF7, F5, and F7 at the left-frontal site. (A) Effects of function word and content word, (B) effects of verb and noun word, and (C) effects of subject and object word

The scalp distribution of P1-, N1-, and P2-peak cross-correlation coefficients are shown as topographies in Figure 16A, Figure 17A, and Figure 18A. Statistical tests (Wilcoxon signed-rank test) were performed at each electrode to test the effect of component exclusion (i.e., function word vs. content word-exclusions, verb vs. noun-exclusions, and subject vs. object-exclusions). Figure 16B, Figure 17B, and Figure 18B show the locations of electrodes that showed significant differences at P1, N1, and P2 peaks between whole-sentence and each exclusion case. The red color indicates that the whole-sentence case shows greater peak amplitude (Wilcoxon signed-rank test, $p<0.05$) whereas the blue color indicates that a component-exclusion case elicited greater P1, N1 or P2 peak amplitude, respectively.

Figure 16 compares the scalp distribution of P1-, N1-, and P2-peak cross-correlation coefficients of the content word- and function word-exclusions cases to the whole sentence case. At N1 peak (199.2 ms), significantly reduced activity was observed in central and left-temporal sites (Wilcoxon signed-rank test, $p<0.05$; Figure 16B top mid panel), which may reflect a weaker evoked response speech perception regions in the left temporal lobe [42]. However, the function word-exclusion case did not exhibit such a difference at N1 peak.

Compared to the case of whole-sentence, the noun-exclusion case elicited significantly reduced N1-peak response also in left temporal and central sites (Figure 17, Wilcoxon signed-rank test, $p<0.05$) whereas the verb-exclusion case did not exhibit such difference at N1 peak.

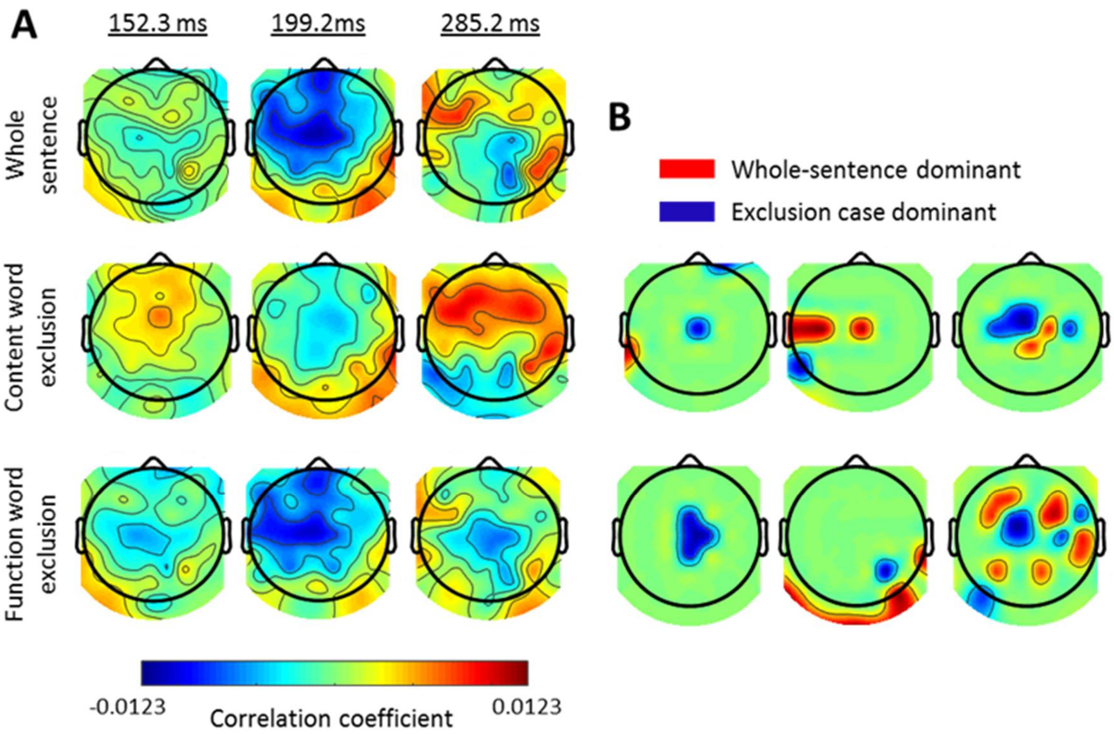

Figure 16: (A) Cross-correlation coefficient function between speech-evoked EEG signals and whole-sentence, content word-exclusion, and function word-exclusion impulse trains at the typical lag of 152.3, 199.2, and 285.2 ms. (B) Dominance map showing whole-sentence versus content word-exclusion and function word-exclusion cases

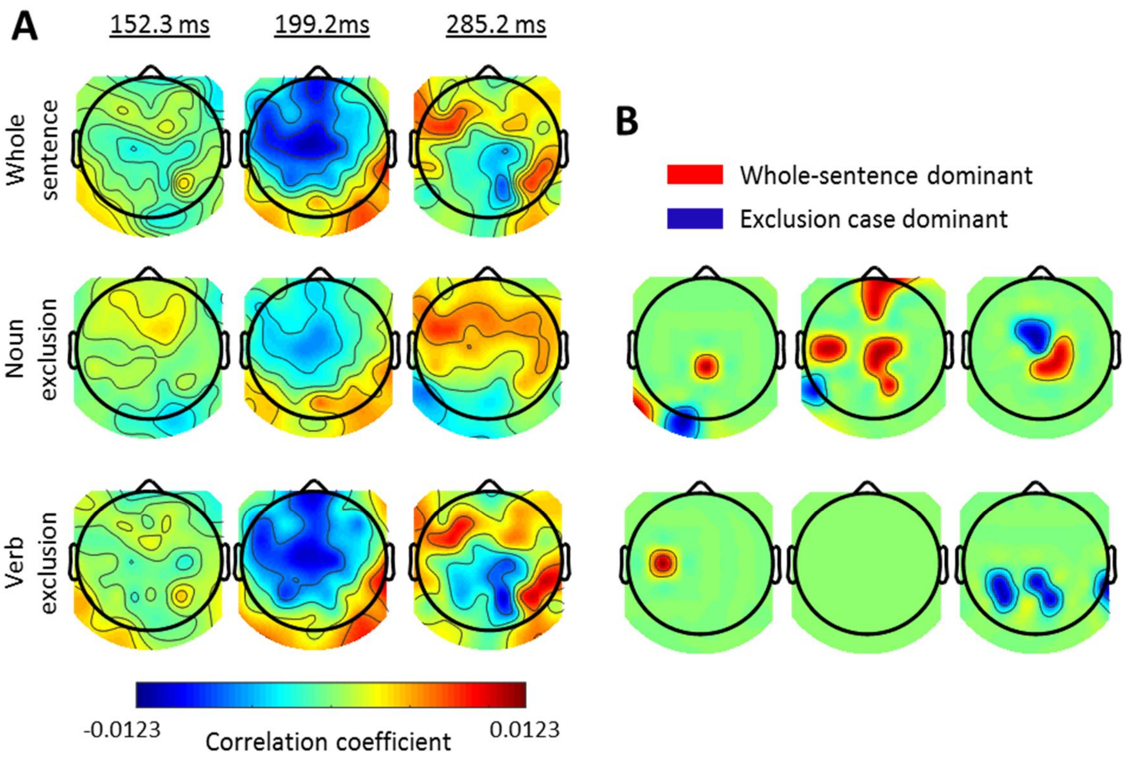

Figure 17: (A) Cross-correlation coefficient function between speech-evoked EEG signals and whole-sentence, noun-exclusion, and verb-exclusion impulse trains at the typical lag of 152.3, 199.2, and 285.2 ms. (B) Dominance map showing whole-sentence versus noun-exclusion and verb exclusion cases

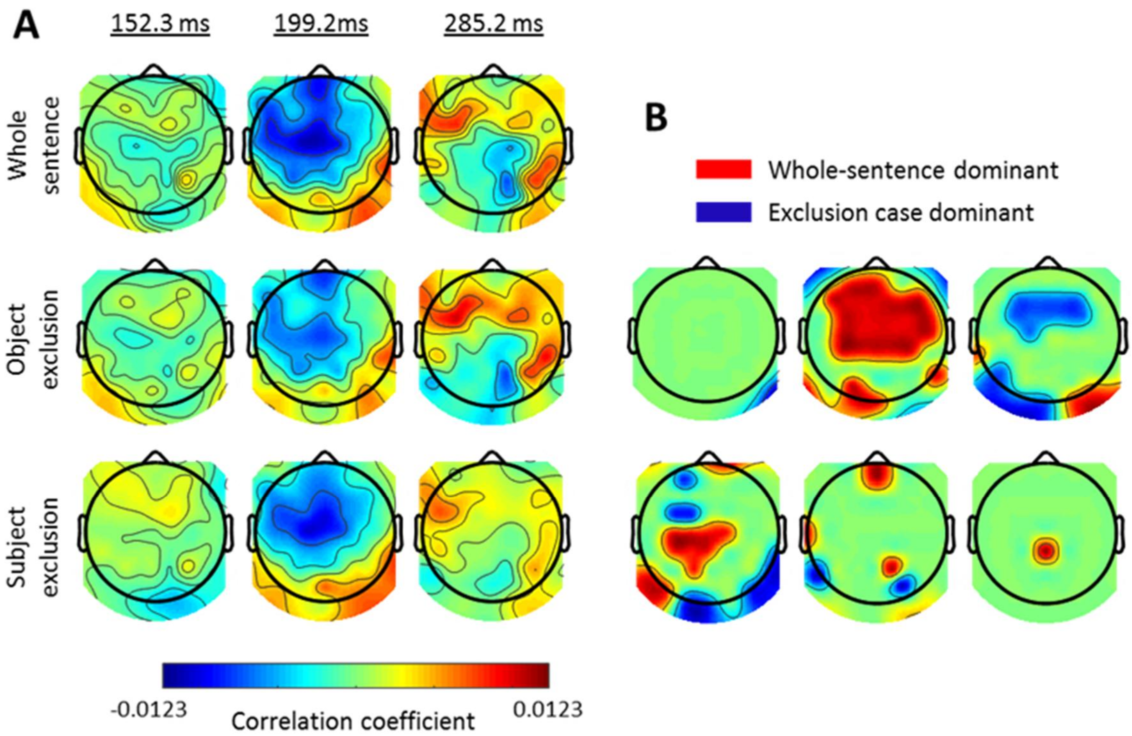

Figure 18: (A) Cross-correlation coefficient function between speech-evoked EEG signals and whole-sentence, object-exclusion, and subject-exclusion impulse trains at the typical lag of 152.3, 199.2, and 285.2 ms. (B) Dominance map showing whole-sentence versus object-exclusion cases

Figure 18 compares the object- and subject-exclusions cases to the whole-sentence case. While the object-exclusion case exhibited significantly reduced correlation coefficients at broad fronto-central and temporal sites at the N1 peak, the subject exclusion case did now show such reduction.

Table 3 summarizes the electrode sites that showed significant differences ($p<0.05$) at N1 peak between each exclusion case and the whole-sentence case in 10-10 standard EEG system. The object-exclusion case exhibited significant reduction of N1 activity at the greatest number of electrode sites (i.e., 31 electrodes out of 64). In contrast, there was none electrode site that showed significant N1 difference in the verb-exclusion case.

4. DISCUSSION AND CONCLUSION

Language processing consists of many stages across multiple time frames. ERP complexes, including N100 (or N1), N400, and P600 have been thought to correspond to different stages of language processing. Sentence comprehension begins with identifying acoustic-phonological events [21, 32] which may be reflected in N1 ERP complex. The current study showed significant reduction of the N1 evoked amplitude when critical linguistic components were omitted from sentences. Given that N1 reflects discrimination of auditory categories [21, 32] while it is strongly modulated by attention [27, 28] the main result of this study may exhibit that the importance of linguistic components is encoded early as the N1 latency. Significant difference at N1 peak time was observed at left-temporal and central regions, consistent with the typical auditory N1 scalp topography, which also strengthens our argument that the preliminary content categorization occurs at phonetic processing stage in early auditory regions (e.g., Heschl's gyrus and planum temporale).

The N1 topography of the whole-sentence case (Figures 17A, 18A and 19A) showed activities in broad electrode positions including temporo-central, frontal, and parietal sites. This is in line with the previous studies that argued speech perception and recognition involve a broad cortical network including frontal-temporo-parietal regions [21, 32]. Our finding adds an argument that such broad activation of cortical network occurs as early as the N1 latency. As observed from topographies in Figures 17B, 18B, and 19B, there is left-hemisphere bias of activities, which is also consistent with previous findings about left hemisphere-dominant language processing [42, 43]. In addition, the object-exclusion case significantly reduced activities in frontal, central sites, which may be consistent with the previous arguments about the contribution of inferior frontal and superior temporal gyri for the assignment of thematic roles [32].

To summarize, the phoneme-based ERP analyses reveal differential importance of linguistic components for sentence comprehension, and such information is encoded early in processing, even during passive listening of sentences. The findings suggest that content words, nouns, and objects are dominant components in sentence comprehension whereas function words, verbs, and subjects, respectively, contribute less to the overall understanding of the sentence.

As the current study is limited to sensor-space analyses, a future study may involve estimation of cortical sources. A future study also include further linguistic components of sentence comprehension such as stress and intonation [2].

REFERENCES

1. Jurafsky, D. and J.H. Martin, *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition*. 2nd ed. Prentice Hall series in artificial intelligence. xxxi, 988 pages.

2. Cutler, A. and D.J. Foss, *On the Role of Sentence Stress in Sentence Processing.* Language and Speech, 1977. **20**(1): p. 1-10.

3. Wang, J., et al., *Identifying thematic roles from neural representations measured by functional magnetic resonance imaging.* Cogn Neuropsychol, 2016. **33**(3-4): p. 257-64.

4. Thothathiri, M., et al., *Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension.* Cognition, 2018. **178**: p. 162-177.

5. Kaan, E., *Event-Related Potentials and Language Processing: A Brief Overview.* Language and Linguistics Compass, 2007: p. 571-591.

6. Beres, A.M., *Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research.* Appl Psychophysiol Biofeedback, 2017. **42**(4): p. 247-255.

7. Alday, P.M., *M/EEG analysis of naturalistic stories: a review from speech to language processing.* Language Cognition and Neuroscience, 2019. **34**(4): p. 457-473.

8. Porbadnigk, A.K., et al., *Single-trial analysis of the neural correlates of speech quality perception.* J Neural Eng, 2013. **10**(5): p. 056003.

9. Bidelman, G.M. and M. Howell, *Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception.* Neuroimage, 2016. **124**(Pt A): p. 581-590.

10. Price, C.N., C. Alain, and G.M. Bidelman, *Auditory-frontal Channeling in alpha and beta Bands is Altered by Age-related Hearing Loss and Relates to Speech Perception in Noise.* Neuroscience, 2019. **423**: p. 18-28.

11. Crosse, M.J., et al., *The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli.* Front Hum Neurosci, 2016. **10**: p. 604.

12. Brodbeck, C., A. Presacco, and J.Z. Simon, *Neural source dynamics of brain responses to continuous stimuli: Speech processing from acoustics to comprehension.* Neuroimage, 2018. **172**: p. 162-174.

13. Viswanathan, V., H.M. Bharadwaj, and B.G. Shinn-Cunningham, *Electroencephalographic Signatures of the Neural Representation of Speech during Selective Attention.* eNeuro, 2019. **6**(5).

14. Horton, C., M. D'Zmura, and R. Srinivasan, *Suppression of competing speech through entrainment of cortical oscillations.* Journal of Neurophysiology, 2013. **109**(12): p. 3082-3093.

15. Kong, Y.Y., A. Mullangi, and N. Ding, *Differential modulation of auditory responses to attended and unattended speech in different listening conditions.* Hear Res, 2014. **316**: p. 73-81.

16. Muralimanohar, R.K., J.M. Kates, and K.H. Arehart, *Using envelope modulation to explain speech intelligibility in the presence of a single reflection.* Journal of the Acoustical Society of America, 2017. **141**(5): p. El482-El487.

17. Di Liberto, G.M., J.A. O'Sullivan, and E.C. Lalor, *Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing.* Curr Biol, 2015. **25**(19): p. 2457-65.

18. Groppe, D.M., et al., *The phonemic restoration effect reveals pre-N400 effect of supportive sentence context in speech perception.* Brain Res, 2010. **1361**: p. 54-66.

19. Mesgarani, N., et al., *Phonetic Feature Encoding in Human Superior Temporal Gyrus.* Science, 2014. **343**(6174): p. 1006-1010.

20. Khalighinejad, B., G.C. da Silva, and N. Mesgarani, *Dynamic Encoding of Acoustic Features in Neural Responses to Continuous Speech.* Journal of Neuroscience, 2017. **37**(8): p. 2176-2185.

21. Friederici, A.D., *Towards a neural basis of auditory sentence processing.* Trends Cogn Sci, 2002. **6**(2): p. 78-84.

22. Ruschemeyer, S.A., et al., *Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian.* Hum Brain Mapp, 2005. **25**(2): p. 266-86.

23. Kutas, M. and S.A. Hillyard, *Event-related brain potentials to grammatical errors and semantic anomalies.* Mem Cognit, 1983. **11**(5): p. 539-50.

24. Kuperberg, G.R., et al., *Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies.* J Cogn Neurosci, 2003. **15**(2): p. 272-93.

25. Frisch, S., A. Hahne, and A.D. Friederici, *Word category and verb--argument structure information in the dynamics of parsing.* Cognition, 2004. **91**(3): p. 191-219.

26. Li, X., et al., *Mental representation of verb meaning: behavioral and electrophysiological evidence.* J Cogn Neurosci, 2006. **18**(10): p. 1774-87.

27. Hillyard, S.A., et al., *Electrical signs of selective attention in the human brain.* Science, 1973. **182**(4108): p. 177-80.

28. Choi, I., et al., *Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography.* Front Hum Neurosci, 2013. **7**: p. 115.

29. Zilles, K. and K. Amunts, *Centenary of Brodmann's map--conception and fate.* Nat Rev Neurosci, 2010. **11**(2): p. 139-45.

30. Gage, N.M. and B.J. Baars, *Decisions, Goals, and Actions*, in *Fundamentals of Cognitive Neuroscience*. 2018. p. 279-319.

31. Friederici, A.D., *White-matter pathways for speech and language processing.* Handb Clin Neurol, 2015. **129**: p. 177-86.

32. Friederici, A.D., *The brain basis of language processing: from structure to function.* Physiol Rev, 2011. **91**(4): p. 1357-92.

33. Keller, S.S., et al., *Broca's area: Nomenclature, anatomy, typology and asymmetry.* Brain and Language, 2009. **109**(1): p. 29-48.

34. Amunts, K., et al., *Broca's region revisited: Cytoarchitecture and intersubject variability.* Journal of Comparative Neurology, 1999. **412**(2): p. 319-341.

35. https://en.wikipedia.org/wiki/Wernicke%27s_area. *Broca's area*.

36. Cohen, M.X., *Analyzing neural time series data : theory and practice*. Issues in clinical and cognitive neuropsychology. 2014, Cambridge, Massachusetts: The MIT Press. xviii, 578 pages, 16 unnumbered pages of plates.

37. https://en.wikipedia.org/wiki/Event-related_potential. *Event-related potential*.

38. *Alphabetical list of part-of-speech tags used in the Penn Treebank Project*. Available from: https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html.

39. Bird, S., E. Klein, and E. Loper, *Natural language processing with Python*. 1st ed. 2009, Beijing ; Cambridge Mass.: O'Reilly. xx, 479 p.

40. Bauer, J. *The Stanford Parser: A statistical parser*. Available from: https://nlp.stanford.edu/software/srparser.html.

41. Paul Boersma, D.W. *Praat: doing phonetics by computer*. Available from: http://www.fon.hum.uva.nl/praat/.

42. Hickok, G. and D. Poeppel, *Opinion - The cortical organization of speech processing.* Nature Reviews Neuroscience, 2007. **8**(5): p. 393-402.

43. Coslett, H.B. and M.F. Schwartz, *The parietal lobe and language.* Handb Clin Neurol, 2018. **151**: p. 365-375.