EEG MEASUREMENT OF NEURAL OSCILLATIONS AT HIGH AND LOW FREQUENCIES ACROSS LANGUAGES

Lauren Stewart

Submitted to the University of Arkansas

Thesis Advisor: Dr. Andrew Bowers

2014

Acknowledgements

It is a pleasure to express the deepest gratitude that I have to all of those who made this thesis possible. First and foremost, to my research mentor, Dr. Bowers. This project would not have been possible without all of your help, support and guidance. From the beginning you have encouraged me to ask questions and my research experience has been invaluable because you have been a part of it. Thank you for giving me your time and allowing me to learn from you.

To my committee members, Dr. Hagstrom and Mr. Aslin. Your willingness to jump on board with this project has meant the world. You have always been a sounding board for me and for that, I thank you.

To the University of Arkansas Honors College. Thank you for the funding that allowed me to further my research. My honors experience has been an incredible one.

Especially to my parents and fiancé Taylor Hudson. Your encouragement and patience have been life giving during this process. You have supported me every step of the way and I am forever grateful for your love.

Finally, to Arsaga's, Mama Carmen's, Onyx and Common Grounds. Thank you for keeping me supplied with fresh caffeine and conversation at all times of the day.

Table of Contents

List of Figures

Abstract

The comprehension and understanding of language depends upon two critical things: first, the acoustic integrity of the linguistic signal that is sent and received as well as the knowledge of the phonology and meaning that is in a given language. However, little is known about how brain rhythms that track these properties involved in comprehension and understanding. A 14 channel telemetric headset measuring electroencephalography (EEG) was used in this study to track the brain's response at both high (30-50Hz) and low frequencies (3-30Hz) while language samples at varying levels of intelligibility were presented to a native Spanish speaker under the following listening conditions: Arabic forward, Arabic reversed, Portuguese forward, Portuguese reversed, Spanish forward and Spanish reversed. Results showed that conditions in which language was comprehended were associated with differences at both high and low frequencies in left and right hemisphere electrodes. A condition in which participants partially understood the intended message (i.e., Portuguese) showed differences from native language listening at high frequencies. Findings are discussed with respect to current theoretical accounts of oscillatory function in speech and language processing along with recent findings.

Introduction

Speech processing is the study of speech signals and the processing methods of these signals in order to understand information. This is the why aspect behind what we hear. In order to understand language, two critical things must be present. First, the linguistic signal that is sent and received must have acoustic integrity. Second, there must be some knowledge of the phonology and meaning that is inherent in a given language. However, little is known about brain rhythms that track both low level acoustic properties and higher level language properties involved in comprehension and understanding. This gap in knowledge is critical as these neural rhythms may be used in assessment and possibly treatment of language disorders. Often a language barrier will mask the fact that an actual language disorder exists and it will instead be written off as trouble learning the language. As such, there is often an overrepresentation of bilingual children in speech therapy programs when what is needed is extra help learning the language. Further work along these lines may in the future help to diagnose brain rhythm differences between those with a true language disorder and those with a language difference.

Speech processing theories predict that multiple brain regions participate in language comprehension, including frontal, temporal and parietal lobes of both hemispheres. More recently, theoretical accounts have suggested that oscillatory assemblies of neurons in these regions share frequency bands. These theories predict that low frequency oscillations track syllable timing that is similar across languages, whereas high-frequency rhythms contribute more to comprehension and understanding of linguistic meaning (Peele and Davis, 2012; Giraud and Poeppel, 2012). Essentially, this means that low frequency waves are showing something that is audibly heard, whereas the high frequency waves are related more closely cognition and understanding. The brain uses different functions to work together to parse and decode speech.

One such way that oscillations are thought to track speech and linguistic signals is through stimulus-brain alignment that occurs when acoustic signals reset oscillations in fronto-temporal networks allowing for the nesting of information at multiple time-scales. For example, phonemes occur at a rate of ~50-60ms, syllables at ~100-250ms, and phrasal boundaries occur between 500-1000ms (Peele and Davis, 2012). These rates correspond with oscillatory cycles of traditional oscillatory bands of the EEG, including gamma (30-70Hz), alpha (8-12Hz), and delta-theta (1-8Hz) bands. Giraud and Poeppel (2012) propose that there is an alignment of auditory and articulatory information, thus resulting in what is eventually the process of deciphering speech into an abstract code (i.e., syllables and phonemes). In addition to low frequency acoustic tracking, recent theoretical proposals have also suggested that higher frequency oscillations track comprehension of language by integrating information at multiple time scales (Peelle and Davis, 2012). However, as yet few studies have investigated changes in time and frequency during the presentation of naturalistic phrase and sentences level stimuli (Peelle and Davis, 2012).

Although few studies have investigated speech signals, one pioneering study using a measure of oscillatory mechanisms demonstrated that the brain processes during speech perception the frequency analyzed was also very clearly broken down into two smaller sub-frequencies; 8-10Hz and 10-12 Hz in the alpha frequency. According to that study, desynchronization of the EEG at alpha frequencies is related to cortical activity (active cognition) whereas the synchronization marks alpha band idling state or deactivation. In that study, participants listened to a speech sample both forwards and backwards and were told to try and memorize the material. When the information was played forward, there was an overall event related desynchronization (ERD) in the alpha band as different mental processes were used to cognitively understand. When the information was played backward and was unintelligible, the greatest event related synchronization (ERS) was observed (Krause, Laine, Lang and Porn, 1997). The implications of this are as follows: when information is actually understood, the neural activity is all over the place and oscillations have huge modulations. When it is unintelligible speech, such as speech played backward, the opposite occurs.

More recent studies have shown that manipulations of intelligibility using time-compression, noise-vocoded speech stimuli, and time-reversed speech have shown that the phase of cortical to stimulus alignment is disrupted when intelligibility is low. In a seminal study, Luo and Poeppel (2007) demonstrated that noisy speech signals are associated with decreases in phase tracking at low-frequencies. However, the same effect was not observed for time-reversed speech in which the acoustic characteristics of the speech were similar (e.g., amplitude modulations and power in the low-frequency spectrum) but comprehension was manipulated (Howard and Poeppel, 2010). Taken together, those studies suggest that low-frequency rhythms are primarily involved in tracking low level acoustic properties of the stimulus. However, as the focus of those studies was on low-frequency rhythms and a measure of phase only, it is unclear how changes in upper frequency rhythms not phase locked to the stimulus respond to low and high frequency.

Despite theoretical predictions to our knowledge, no studies have investigated time-frequency changes to native and non-native languages at the sentence level across frequencies. Given current theoretical accounts acoustic processing and comprehension should modulate low and high frequencies differentially. Further, EEG studies of oscillatory activity typically use stationary equipment costing in the tens of thousands, making the use of such equipment impractical in some environments. For example, although many speech and hearing clinics have acquired such equipment, most still do not have access to multichannel EEG. As such, the aims of the current study are to investigate changes in high and low frequencies by manipulating level of comprehension relative to a low level acoustic baseline in which speech samples are acoustically reversed using a low-cost telemetric EEG cap and signal acquisition software on a lap-top computer.

Given the predictions of theoretical accounts, low frequencies (below 30Hz) would be expected to show changes primarily related to acoustic processing. By contrast, high frequency oscillations would be expected to show changes related more closely to language comprehension. Whereas reversed stimuli still contain the spectral and amplitude modulations associated with speech (Peele and Davis, 2012), reversed stimuli may be used as an adequate baseline. Thus, a comparison between two signals containing intelligible (i.e. comprehensible) speech presented in a subjects native language and acoustically similar unintelligible speech would be expected to show differences in the high-frequency rhythms. Further, a language closely related to a participant's native language in which comprehension is sparse would be expected to show differences at high frequencies when compared to the participant's native language.

Methods

Design

To investigate the proposed hypotheses, this study employed a design in which language samples were presented at different levels of intelligibility to a native Spanish speaker under three different conditions: 1) native language (Spanish); 2) a closely related language (Portuguese); 3) entirely unrelated language (Arabic). Arabic was chosen because it is entirely different from Spanish. Portuguese was chosen because it is more similar to Spanish than Arabic and is more of a blend of Spanish and French. Each condition was also compared to a reversed sample to provide a baseline for that language, giving a total of 6 listening conditions. It was expected that low-frequency rhythms would be associated with forward and reversed speech. However, significant differences and modulations were expected in high-frequency rhythms associated with greater intelligibility.

Participants

Seven native Spanish speakers (2 males and 5 females) who had not been diagnosed with any type of speech or language disorder were used for this study. All participants were between the ages of 18 and 24 and had been recommended to participate in this study. Prior to the experiment, written informed consent approved by the University of Arkansas Institutional Review Board was obtained for all participants and compensation in the form of a Walmart gift card was given for participation.

Stimuli

Speech stimuli consisting of a randomly chosen two-sentence sample was taken from a middle-school level reading passage (Flesch-Kincaid grade level; 8.3). The language sample was recorded from three native speakers of the following languages: Spanish, Portuguese, and Arabic. Garage Band software on an Apple Macbook (2 Ghz processor) was used to record the language samples. All samples were root-mean square normalized (RMS) using Adobe Sound Edit on a Dell desktop computer (3.4Ghz) so that all samples were presented at the same amplitude level. The use of native speakers and their own translation of the reading material allowed an authentic recording in their native language, eliminating translation errors and gaps in meaning. Each translator was a female so that tone and pitch remained as close in range as possible across samples. Each sample was edited so that it was 13 seconds in length. Samples were edited such that the end of the sample coincided with a word boundary. Spoken language samples were presented in the following conditions: 1) Arabic forward (Arb_for), Arabic reversed (Arb_rev), Portuguese forward (Port_for), Portuguese reversed (Port_rev), Spanish forward (Span_for) and Spanish reversed (Span_rev).

Procedure

Stimuli were presented using BCI 2000 stimulus presentation software on a Hewlett Packard laptop computer. Each participant was placed in the sound booth in Room 173 in Epley Center where they were fitted with headphones, and seated in front of a 10 inch monitor for stimulus presentation. Data was acquired at a sampling rate of 128Hz. The Emotiv 1.0 EEG headset was placed on the head according to the 10/20 system of electrode placement. Data collection was monitored on a separate laptop computer. Once the headset was in place, the participant was told to listen and attend to the speech sample as questions about the sample were to be asked afterward. Stimuli were presented in blocks of 20 trials after which the participant was allowed to rest. Special care was taken with frontal and temporal channels known to contribute to non-neural noise from eye-blinks, horizontal eye-movements, and temporal muscle movement (e.g. jaw clenching). To reduce noise due to movement, Independent Component Analysis (ICA) employing the *binica* algorithm was applied to each sample. ICA is a robust method of noise removal and has been shown in previous studies to eliminate repeated non-neural sources (e.g. eye-blinks) (Delorme and Makeig, 2004).

Analysis

Noise free trials were averaged and changes in the spectrum from 3-50Hz were analyzed using EEGLAB script written in a MatLab environment. Event-related desynchronizations (ERD; power decrease) and event-related synchronizations (ERS; power increase) were assessed using a time-frequency map. First, each time-frequency map was computed relative to a silent recording interval for each condition across participants. Second, condition differences were tested using a repeated measures ANOVA design using the STUDY command structure in EEGLAB. Planned paired comparisons were used to determine condition effects related to experimental hypotheses using *t*-tests across the entire time-frequency matrix. As temporal lobe regions are thought to be critical for receptive language processing, analysis focused on left (T7) and right (T8) electrodes locations. To show the overall spatial distribution of power changes, maps of the overall-scalp distribution were shown for high and low frequency bands of interest (see Figure 3).

Results

Significant p-values for each time-frequency bin over left (Figure 1A) and right (Figure 1B) temporal electrodes (T7 and T8) from 3Hz to 50Hz are shown in Figure 1. To test condition effects, a one way repeated measures ANOVA design with the factor condition (1x6) was employed. For the omnibus test, significant differences (uncorrected) in time-frequency values were found in the delta-theta range (3-7Hz), alpha/beta range (8-30Hz) and in the gamma range (30-50Hz). The highest p-values were noted in the alpha/beta and gamma ranges (dark red).

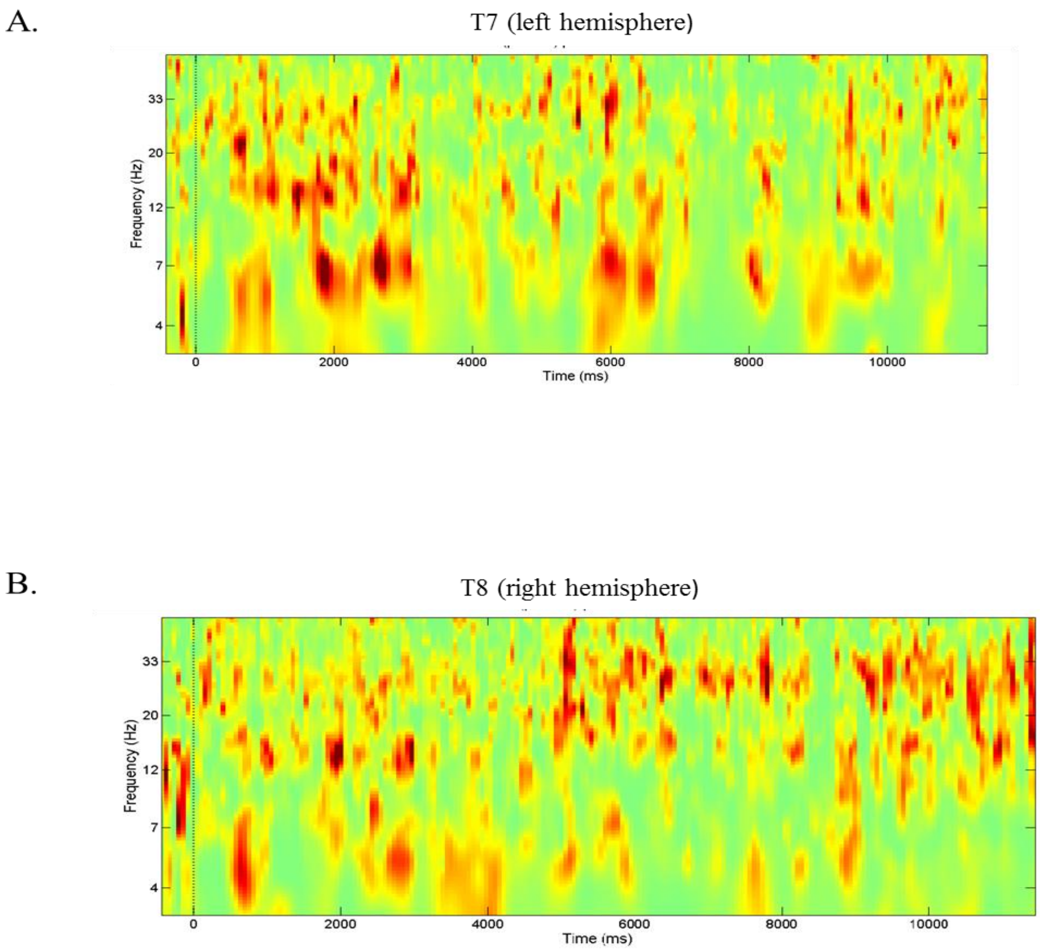

Figure 1. Time frequency for temporal electrodes

Mean time-frequency values for planned comparisons from 3-50Hz are shown in Figure 2. First, to assess whether languages played forward were different from the same language sample played backward, a paired *t*-test was used to compare each language to its reversed counterpart. A comparison between Arb_for and Arb_rev (Figure 2A) showed no significant differences, with significant power enhancements (ERS) across frequencies relative to silent baseline. A comparison of the Port_for and Port_rev conditions (Figure 2B) showed significant ERD in the gamma band (30-50Hz) for the Port_for condition relative to the Port_rev condition.

Finally, a comparison between the Span_for and Span_rev (Figure 2C) conditions showed a significant difference across frequencies (3-50Hz).

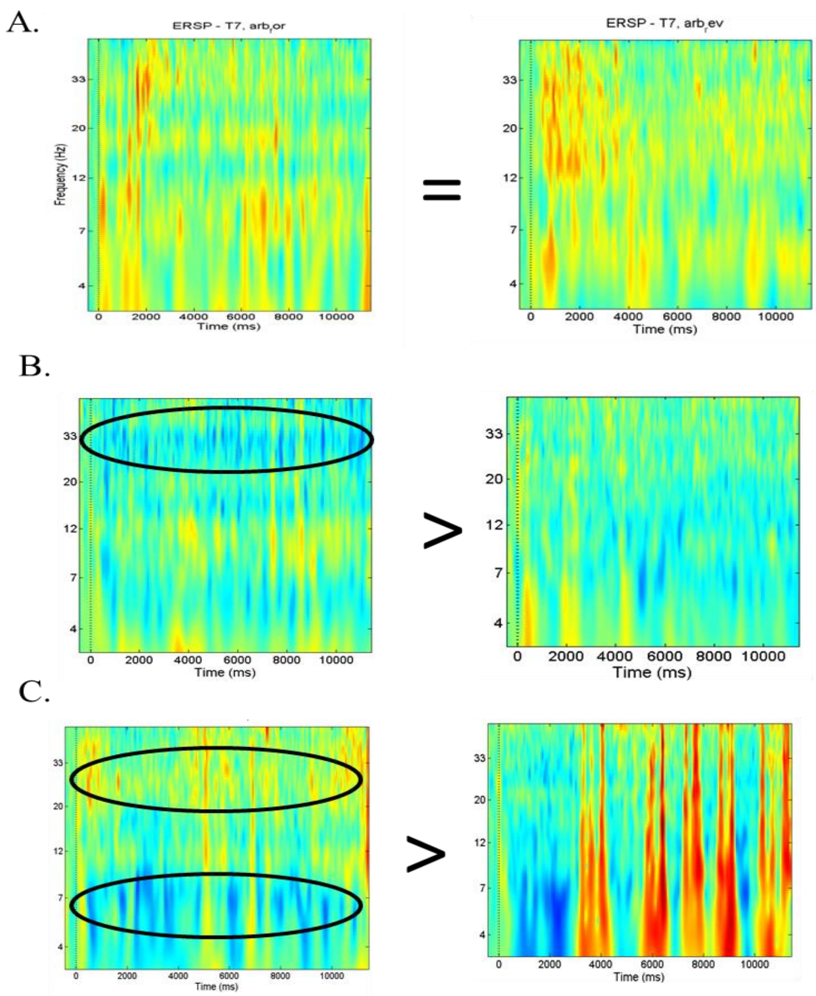

Figure 2. Mean time-frequency comparisons for 3-50 Hz for Arabic (A), Portuguese (B) and Spanish (C)

As one of the initial hypotheses was that a language similar to the participants' native language would show intermediate effects, the Port_for and Span_for were compared using a *t*-test design (Figure 3A). Results indicated significant differences (uncorrected) at alpha/beta and gamma frequencies. Relative to Portuguese, the participants' native language was associated with ERS at gamma frequencies and ERD at alpha frequencies. The scalp distribution across electrodes is shown in Figure 3B at alpha/beta (7-15Hz) and midgamma (30-50Hz) frequencies.

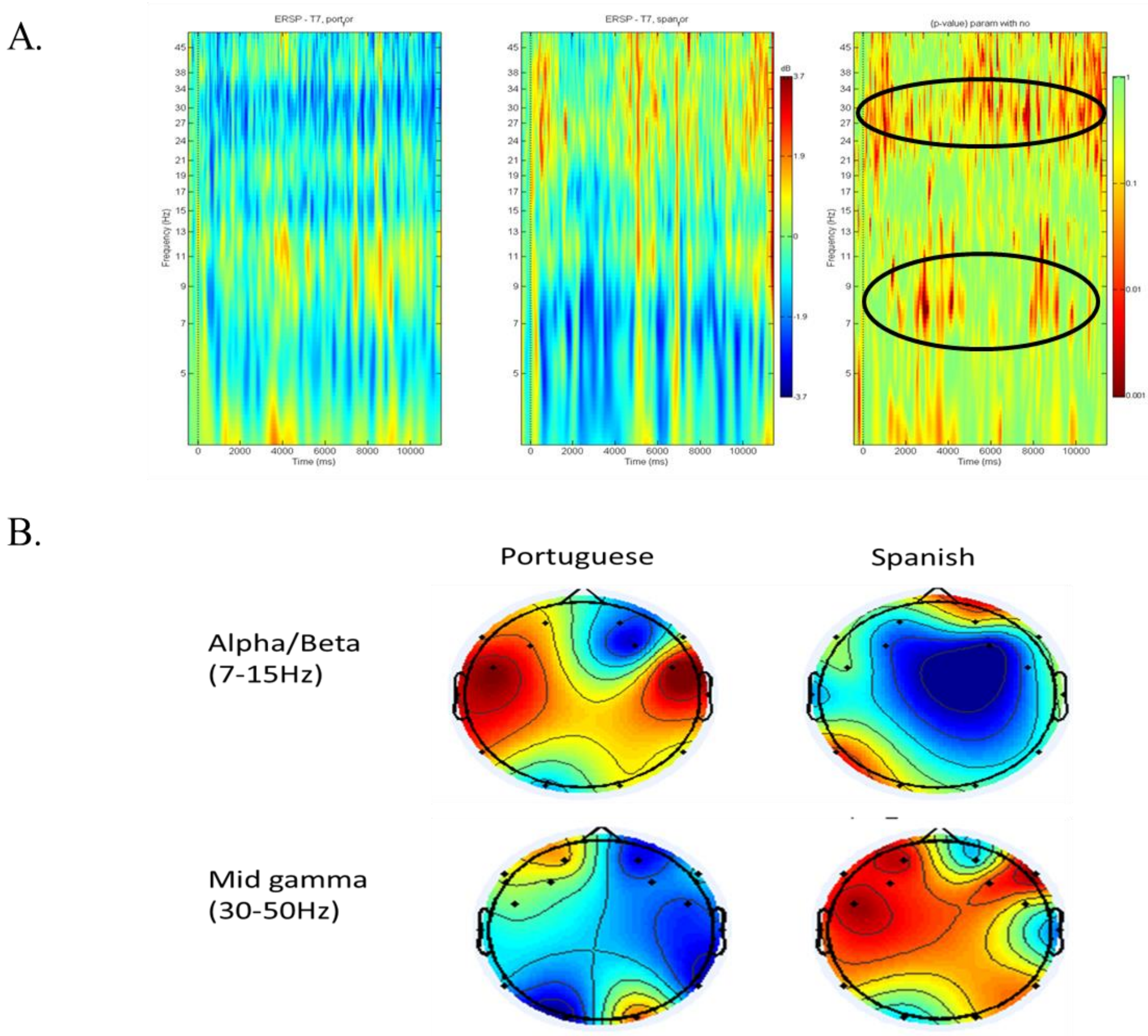

Figure 3. Direct comparison of Portuguese and Spanish using t-test (A) and scalp distribution comparison of Portuguese and Spanish (B)

Discussion

The purpose of this study was to use a portable EEG apparatus to evaluate whether or not it could be used to aid in determining how a native Spanish speaker responds to acoustically reversed, non-native language, and native language as well as to analyze the differences that would occur between conditions. According to initial hypotheses, an examination of both high and low frequencies was expected to yield information about mean event-related synchronization and desynchronization (ERS and ERD) that would in turn give a picture of what frequencies across electrode locations are related to language comprehension relative to acoustic processing without comprehension. Generally, the results suggest that a 14 channel telemetric cap is sufficiently sensitive to capture differences in language comprehension relative to acoustic processing. First, there were no differences in processing for Arabic played forward and backward which would suggest that acoustic processing in the absence of understanding is associated with little change in either high or low frequencies. Second, consistent with initial expectations, for stimuli in which the participants' native language was comprehensible were associated with significant differences in upper frequencies (i.e., mid gamma) thought to be more closely related to comprehension along with segmental processing of the signal (i.e., syllable and phoneme units). In the discussion following, findings will be discussed relative to current theoretical accounts and the few studies in which ERD and ERS have been measured during the presentation of naturalistic linguistic stimuli.

In the theoretical account proposed by Giraud and Poeppel (2012) they suggested that throughout the 1-72 Hz frequency range observed that speech modulations were not tracked evenly by cortical activity. Instead, it was the theta and gamma domains that were seen to have shown the most cortical activity. It was also evident that when envelope tracking failed, the ability to intelligibly understand speech was compromised (Giraud and Poeppel 2012). In the instance of Arabic forward and reversed, an entirely unfamiliar language, the participants had no prior knowledge and therefore were unable to code what they were hearing. There was no alignment and therefore no speech processing could take place. By comparing differences in the alpha and gamma bands for Arabic, Portuguese and Spanish (all forward and reverse), there is a particular difference when comparing the results of Spanish to that of Arabic. The comparison of Arabic conditions results in almost no difference relative to the reversed baseline. The language was completely unknown so hearing it forward or reversed would have no significant bearing on the results. Peelle and Davis (2012) suggest that it is the low frequency oscillations that provide a framework for rhythm. This fits with findings in all low frequencies that sound is detected and the brain is able to pick up on a rhythm when played both forward and backward. However, because there is no intelligibility for incomprehensible or unfamiliar samples, there is not a significant change in the higher frequencies.

While the results of this study are broadly in line with recent theoretical accounts, little time-frequency data has been recorded in response to naturalistic sentence level stimuli. The Krause et al. (1997) study showed that the alpha frequency band that showed a response to linguistic activity and that this is where the most cognition, understanding, or memory for an utterance takes place. When participants listened to a particular Finnish text forward and reversed, it was the text forward stimulus that elicited the greater ERD in the alpha band (8-12Hz). This matches closely with the participants' responses to their native language played forward (i.e., Span_for). The Krause et al., (1997) study suggests that some response in both alpha frequency bands, however the lower frequency bands were hypothesized to have only correlated with sustained attention. Given the common interpretation that ERD in the alpha band is an index of attention and memory, stronger alpha ERD in the current study for Spanish_for when compared to Port_for could be interpreted as cognitive processing related to understanding and retrieving previous experience with the participants' native language.

One of the aims of this study was to see if Portuguese would fall somewhere in the middle between Spanish and Arabic. Findings suggested that Portugese_for condition was associated with ERD in the gamma band as opposed to enhancement relative to the Port_rev condition. Portuguese also showed a limited amount of ERD in the gamma band relative to the Span_for condition. Although it is unclear why ERD in the gamma band was observed for the intermediate language condition, in general ERD is interpreted as a disinhibition of ongoing cortical activity reflecting increased stimulus processing. As such, one potential explanation for ERD in the gamma range for the Port_for condition is increased processing related to incomplete understanding of the communicative message. More specifically, in perceptual contexts, it has been proposed that the gamma band disinhibition may be related to processing increased error signals at phonemic time-scales (Arnal and Giruad, 2012), suggesting that gamma band ERD here may have been related to continuous errors in processing phonological inventories that differ from previous experience with a closely related native language (i.e., Spanish).

A more recent article was discovered after data collection and analysis was complete. A surprisingly similar study to this one was completed in 2012 by Pena and Melloni who had native and non-native speakers of Italian and Spanish listen to language samples in Italian, Spanish and Japanese both forward and reversed. That study found that participants displayed an increase in gamma band power only when they heard their native language played forward relative to other languages played forward. The gamma power increase shown in the native language was interpreted as a temporal binding mechanism reflecting the integration of acoustic, paralinguistic, and syntactic/semantic aspects of the signal (Pena and Melloni 2012). Alpha band suppression was also noted but was not significantly different from an intermediate language (e.g., Italian for Spanish speakers) and was interpreted as a general measure of attention to the signal. By contrast, while the current study implicates changes in gamma ERS as an important marker of linguistic comprehension, alpha band ERD was also different from the intermediate language (i.e., Portugese). Consistent with Krause et al., (1997) alpha ERD may also be a sensitively marker of comprehension, perhaps more closely related to general cognitive function (e.g., attention/working memory). In summary, suppression in the alpha band and enhancement in the gamma bands appears to reflect the integration of the multi-time varying aspects of a spoken speech signal into abstract units for comparison with previous linguistic experience.

Limitations

Several limitations are applicable to the current design. First, while the Emotiv headset is inexpensive and highly accessible, with only 14 channels little spatial information can be recovered from the signal. Much like speech, the EEG is a highly complex time-varying signal that also significantly varies in space as it varies in time. Further, the recovery of spatial information may be critical to interpretations of the signal and how it is processed in distributed networks across the fronto-temporal network. Second, one confound for this study is the fact that data presentation order was not randomized, instead it was presented in the same order each time. Thus, the results from linguistic manipulations might have interacted with an effect for order. However, in that case alpha ERD, would be expected to diminish over the course of the experiment rather than to increase in the Span_for condition (presented in the next to last order). Another limitation for this study is that only 7 participants were recruited due to time limitations for study completion. Greater study power could be achieved with a larger number of participants and more stringent statistical tests (i.e., with corrections for multiple comparisons) could be used. A longer and more varied sample might also be beneficial in keeping the attention of the participants and in determining what aspects of the signal mediate responses. Despite these limitations, this study demonstrates the sensitivity of a highly accessible and portable EEG apparatus for measuring time-frequency differences related to comprehensible language processing.

Future Directions

While this data was being collected, the only known similar study was the 1997 study from Krauss. However, the Italian study by Pena and Melloni study was found after data analysis had taken place. The findings are quite similar; alpha and gamma differences reflection integration However, this study was done much less expensively. Rather than thousands, only a few hundred dollars were spent on equipment to record and analyze neural oscillations. The EEG equipment in this study is portable and could easily be used in different locations.

Ideally this study would be transferrable to the clinic. This study is reproducible and more cost efficient than traditional approaches to acquiring EEG. In summary, this study contributes to a better understanding of the brain's response to familiar and unfamiliar language, both comprehensible and incomprehensible. With the difference in neural rhythms so easily at hand, there is potential to bring this to a clinic as a diagnostic tool. In a world where bilingualism is growing rapidly, diagnoses of difference versus disorder are often being overlooked. Because of the masking effect of a second language, the learner will often be given a false positive or a false negative. With more work, the current approach might be used to develop a profile of time-frequency markers similar to those used in quantitative EEG approaches to the differential diagnosis of psychological disorders. A future study might involve the use of Marshallese children in the Northwest Arkansas area both with and without a known language disorder as participants. The technology would be used to play language samples in English and Marshallese to track time-frequency responses in both populations. Based on theoretical accounts and the current results differences in the midgamma and alpha bands for those with a disorder would be expected.

References

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Delorme, A and Makeig, S. (2014 March 15). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21

Giraud, A., Poeppell, D. (2012 April). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15* (4), 511-517

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, *104*(5), 2500.

Johnston, S., Mennen, I., Stansfield, J. (2005). Speech and language therapy services for bilingual children in england and scotland: a tale of three cities. ISB4: Proceedings of the 4th International Symposium on Bilingualism

Krause, M., Laine, M., Lang, A., Porn, B. (1997 February 4). Relative alpha desynchronization and synchronization during speech perception. *Cognitive Brain Research, 5,* 295-299

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

Peelle, J. and Davis, M. (2012 September 6). Neural oscillations carry speech rhythm through to comprehension. *Frontiers Research Foundation, 3,* 1-17

Pena, M. and Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience, 24*(5), 1149-64