USING COGNITIVE NEUROSCIENCE TO UNDERSTAND LEARNING MECHANISMS: EVIDENCE FROM PHONOLOGICAL PROCESSING

by

Enes Avcu

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Summer 2019

USING COGNITIVE NEUROSCIENCE TO UNDERSTAND LEARNING MECHANISMS: EVIDENCE FROM PHONOLOGICAL PROCESSING

by

Enes Avcu

Approved: ____
Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics and Cognitive Science

Approved: ____
John A. Pelesko, Ph. D.
Dean of the College of Arts and Sciences

Approved: ____
Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate & Professional Education and Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Arild Hestvik, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Jeffrey Heinz, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
William Idsardi, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Zhenghan Qi, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Edward Wlotko, Ph.D.
Member of dissertation committee

What is essential is invisible to the eye.
—The Little Prince, Antoine De Saint-Exupéry, 1943

To Canan, Kerem and Zeynep

ACKNOWLEDGMENTS

I am so grateful for so many people who encouraged, helped, inspired, and motivated me as I completed this task—it was not easy enough to be done on my own.

I will first acknowledge that my wife Canan has provided me with a huge amount of support during my doctorate years. While I can recognize certain sacrifices, she made and the amount of faith she put in my abilities, I do not believe that I can fully understand how difficult it is to be on the other side of life as a Ph.D. student. Nonetheless, I offer her an assuredly heartfelt "Thank you!", along with the hope that our adventure in Delaware was both fun and worthwhile. Our son Kerem and our daughter Zeynep also brought smiles to my face on the days I thought nothing could break through thick fog.

I owe a great deal of thanks to my parents, and to my brothers and to my sister who have always supported me and have rarely questioned my lofty goals. Thank you all, from the bottom of my heart.

Like many others in this field, I was initially drawn to linguistics by the enthusiasm of those teaching my introductory courses. I, therefore, want to recognize two instructors from my time at Hacettepe University: Aydogan Yanilmaz and Zeynep Doyuran who were among the best teachers I have ever had. I also wish to thank, Cem Bozşahin, Annette Hohenberger and Deniz Zeyrek for the guidance they provided me throughout my master years at Middle East Technical University. This thesis would not have been possible without their support.

I had the pleasure of working closely with a number of UD faculty members who have influenced my research and my perception of what it means to be a scholar.

In addition to those who taught my graduate courses, I would like to mention Satoshi Tomioka and Ben Bruening, who provided excellent demonstrations of what it will take to succeed early in my career.

I gratefully acknowledge the help of the students of the University of Delaware who participated as subjects in this project, and the collaboration of my classmates from the Department of Linguistics and Cognitive Science, in particular, Jinwoo Jo, without him, I could not have survived the first year of Ph.D.; and Ryan Rhodes and Chao Han. I also appreciate the help that was given to me by Kelsey Mulford during the participant recruiting process.

It is difficult to overstate the impression that Jeffrey Heinz has made on me. I have observed and interacted with him in his role as a teacher, as a researcher, as a graduate advisor, and as a member of my committee. I continue to be struck by his respect for others, his general curiosity, his brilliance, and his focus on the importance of everyone's work-life balance.

Inevitably, however, the person whose influence is most apparent in my work is Arild Hestvik. He provided the facilities and equipment for all of my experiments, he stressed the importance of methodological detail and statistical rigor in my research, and his role in meetings was often to ensure that my tasks could actually be achieved on time. I am left wondering how anyone completes a Ph.D. without Arild on their committee. I could not have asked for a better advisor and I would not have been successful without him. Despite spending hundreds of hours together discussing research, I cannot pinpoint why exactly we work so well together. I can only hope that I will be able to exemplify similar qualities as a person and a psycholinguist.

I have also benefited from the input and advice of several other people. In particular, discussion and correspondence with William Idsardi who was always accessible and never said NO! for a meeting. Moreover, Edward Wlotko, Zhenghan Qi, Roberta Golinkoff and Kaja Jasinska have provided more insight than they may realize, and my research is significantly better because of them.

Finally, I point out that my dissertation research was supported financially by the Turkish Armed Forces, University of Delaware Dissertation Fellowship and the Linguistics and Cognitive Science Department.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation studies different learning mechanisms of phonological processing by conducting behavioral and neurophysiological experiments in the artificial grammar learning paradigm. The main goal is to identify the phonological computations that give rise to the complex combinatorics underlying human languages by providing new knowledge about whether linguistic constraints that are learned in laboratory situations are directly “channeled” into incremental, real-time phonological predictive processing. It also explores whether the learning outcome depends on the type of learning mechanisms (domain-specific vs. domain-general) and the computational complexity of the patterns. It achieves this by examining the nature of the neural commitment produced by language exposure via electroencephalogram (EEG).

The specific patterns that were tested in this dissertation are well-motivated from the perspective of formal language theory and theoretical linguistics. Processing of these phonological patterns must ultimately be integrated into the complete cognitive neuroscience of language that links abstract mathematical grammars to real-time theories of word processing and ultimately to biologically plausible neural computations. The dissertation, therefore, contributes to a growing body of research that highlights the nature of acquired linguistic knowledge at the neural level that behavioral measures cannot address. It provides new insight into how artificial language learning causes acquisition mechanisms to incorporate linguistic rules acquired in the laboratory.

This dissertation’s contribution to knowledge is threefold. Firstly, it has been shown that human phonological learning mechanisms are restricted mainly by domain-specific computational constraints. Secondly, specific phonological rules that are

acquired in laboratory situations are directly channeled into incremental, real-time phonological predictive processing evidenced at the millisecond level by ERP components. Thirdly, domain-specific learning mechanisms associated with implicit learning and domain-general learning mechanisms associated with explicit learning converge on similar knowledge states, but with different underlying neural mechanisms.

The hope is that these results will illustrate how formal language theory can be integrated with the cognitive neuroscience of language, and how experimental techniques from experimental psychology and cognitive neuroscience can reveal new insights into the critical properties of human language such as the complex interplay between domain-specific and domain-general components of cognition.

Chapter 1

INTRODUCTION

This dissertation explores the neurophysiological correlates of phonotactic pattern learning. In particular, I address the question of how generalizations acquired in laboratory settings are cognitively encoded. I mainly focus on non-adjacent phonotactic dependencies and the types of locality relations that are observed in such patterns. The primary goal is to provide new knowledge about whether linguistic constraints that are learned in laboratory situations are directly “channeled” into incremental, real-time phonological predictive processing, and whether this process depends on the type of learning mechanisms (domain-specific vs. domain-general) and the computational complexity of the patterns. I achieve this by examining the nature of the neural commitment produced by language exposure via electroencephalogram (EEG) in a series of artificial grammar learning experiments.

As described by Carson (1988), the phonological parser is a processing mechanism that uses a phonotactic network to parse the phonetic sequence into phonemes and syllables. If a pattern has been learned, this means that the phonological parser has incorporated the rule into a mental grammar, which entails that the neural correlates of the decision processes that rely on the rule should be observable as part of real-time parsing mechanisms. Given that phonological parsing is incremental (Ferreira & Swets, 2002; Kempen & Hoenkamp, 1987; Levelt & Meyer, 2000; Levelt, Roelofs, & Meyer, 1999; Wheeldon, Meyer, & Smith, 2005), I expect laboratory-induced learning of phonotactic rules to be reflected in real-time predictions during incremental word parsing.

Exposure to statistical regularities in language stimuli automatically results in the formation of language-specific neural networks in the brain (Aaltonen et al., 2008). The Native Language Neural Commitment (NLNC) hypothesis (Kuhl et al., 2008) states that language exposure (to L1) produces a "neural commitment" that affects future processing and learning (of L2). That is, initial exposure to a language stimulus causes physical changes in neural tissue (committing neural circuitry) which in turn reflects properties of the language input. In terms of neurophysiology, this implies that if an abstract rule is learned after exposure to repeated stimuli, this learning should translate into observable neural prediction mechanisms. Having learned the rule means that the brain comes to expect certain patterns in the input, and violations of the pattern should result in an expectation violation response (Bubic, von Cramon, & Schubotz, 2010). In other words, acquired rules should lead to neurophysiological reflexes in response to violations of those rules. What is less certain is whether such neurophysiological responses will only be detectable after years of "neural commitment" during natural language acquisition and development, or whether they also can be observed after very brief exposure in laboratory-based learning contexts in adults.

Pattern extraction, whether in the laboratory or as part of human language acquisition, is essentially implicit – meaning that the learning does not come with conscious awareness of the rules. The term “implicit learning” in Reber (1967)’s sense refers to the unintentional or unaware knowledge acquisition of complex patterns. In contrast, the term “explicit learning” refers to the acquisition of conscious or intentional knowledge (Rebuschat, 2015). Moreton & Pertsova (2016) classify implicit learning as effortless, unconscious, and gradual, whereas explicit learning is effortful, conscious, and abrupt. While implicit learning requires only positive examples without feedback, explicit learning is facilitated by training with positive and negative feedback. Another distinction regarding implicit vs. explicit learning mechanisms was put forward by Moreton, Pater, & Pertsova (2017): cue-based (implicit) learning refers to the learning of a rule using cues to make an inference; whereas the rule-based (explicit) learning refers to learning where the learner consciously uses a rule to make an inference. Consequently, it is evident that these two learning processes have different properties that are triggered by different conditions (Maddox & Ashby, 2004). However, less is known about how these two learning systems are encoded at the neural level: Is there a difference in neural coding between implicit and explicit learning?

Language patterns have many levels that pose different learning challenges to the learner according to their inherent difficulty. The Chomsky Hierarchy (Chomsky, 1956), a categorization of formal grammars, is one representation of such inherent difficulty. It divides all logically possible patterns into nested regions of complexity. While sentence well-formedness (syntactic) patterns inhabit a range of regions in the hierarchy from regular to context-sensitive (Kobele, 2006), word well-formedness (phonology) patterns are in the regular region (Kaplan & Kay, 1994). Heinz (2010) further shows that phonotactic patterns in natural languages inhabit proper subsets within the regular region. The Subregular Hypothesis (Heinz, 2010) claims that humans' phonological pattern detectors are constrained. In particular, there is a learnability distinction between the different phonological patterns that reside in different regions. I will test this specific hypothesis about the computational complexity of phonotactic patterns and their learnability. One prediction of this theory is that, in terms of computational complexity, a local rule like nasal spreading is computationally easier than a non-local rule like vowel harmony; and a local rule like Tagalog partial reduplication is easier than another local rule like Rotuman deletion because the latter includes a more complex local relation between the segments (Chandlee, 2014). Empirical evidence for these learnability predictions was previously obtained using a behavioral task (Lai, 2012, 2015). What I propose is to assess these learnability distinctions with attentive (or involuntary, depending on the design of the experiment) brain response measures.

The central questions posed by this dissertation are as follows:

- Will neural commitment caused by phonotactic pattern learning be detectable after very brief exposure in a laboratory setting?
- Do different types of exposure (implicit vs. explicit) lead to different neural commitments?

- Are computational complexity constraints on phonotactic learning of varying complexity encoded differently at the neural level? Specifically, can proposed computational constraints on phonotactic learning be assessed by using event-related brain responses to rule violations?

This dissertation, therefore, contributes to a growing body of research that highlights the neural nature of acquired linguistic knowledge that behavioral measures may not address. In particular, this research would provide new insight into how artificial language learning causes acquisition mechanisms to incorporate linguistic rules acquired in the laboratory.

A series of experiments was conducted to answer the above questions. The first experiment (Experiment 1) was designed to compare behaviorally the relative learnability of two non-local harmony patterns that differ typologically (attested vs. unattested) and computationally (simple vs. complex). Lai (2015) provided the initial laboratory evidence for this learnability distinction; the aim of Experiment 1 was to replicate and extend these findings by using a different experimental paradigm (oddball task) and a different measure of learning (sensitivity index, d'). The results of Experiment 1 show that there are significant learnability distinctions between these two phonological patterns, but the unattested patterns also show signs of relative learnability. These results substantiate that the Subregular Hypothesis is active during real-time phonological parsing and to a significant extent constrains the learnability of specific phonotactic patterns. However, this study’s real contribution to the literature lies behind the fact that the unattested patterns’ learnability implies the use of different learning strategies. In particular, laboratory participants brought domain-general learning to bear, which made learning the unattested patterns possible. However, the attested and the computationally simpler pattern was much more readily learned because it was detectable by specific phonological learning mechanisms.

The second series of experiments were preliminary in the sense that they were exploratory and designed to examine the neurophysiological signatures of phonotactic learning. The aim of these experiments was to find an appropriate design which will allow us to measure the neural commitment or elicit brain responses reflecting phonotactic rule learning. First, an artificial grammar learning experiment was run to test whether violations of a long-distance sibilant harmony patterns trigger a prediction error response, namely the Mismatch Negativity (MMN) (Experiment 2). The MMN, which is one of the auditory event-related potentials (ERP), reflects a change detection response in the brain (Näätänen, Paavilainen, Rinne, & Alho, 2007). Participants had a training session (repeating grammatical exemplars without any explicit instruction) and a test session (while recording brain responses with EEG, participants listened to words in an auditory oddball paradigm and responded behaviorally to words violating the pattern). The results show no MMN in response to violations of such patterns. Some post-hoc explanations were discussed for the lack of brain response. One possible explanation was that we might just not be able to measure it because the MMN is not an appropriate tool to measure this kind of learning. Another explanation was related to MMN's not having access to higher-order regularities. This possibility was also reported in the literature which was discussed towards the end of the related chapter. Next, another artificial grammar learning experiment testing the same pattern with a different test design (Experiment 3) was run. Participants had a training session (repeating grammatical exemplars without any explicit instruction) and a test session (while recording brain responses with EEG, participants rated how likely it is that each word is part of the language they were learning). I hypothesized a higher amplitude Late Positive Component (LPC) to novel words that violated the learned phonotactic constraint than novel words that satisfied it. The LPC can be taken as an index of detection of an anomaly in rule-governed sequences (Núñez-Peña & Honrubia-Serrano, 2004). The behavioral results showed that participants were able to make a distinction between novel-fit and novel-violate words. However, the LPC was not elicited in response to novel words violating the phonotactic pattern. In fact, it was entirely absent from the data. These results imply that the LPC may not be indexing violations of long-distance phonotactic rules, but it may depend instead on local sequencing. Nevertheless, it is hard to make definite conclusions from the null results.

The fourth experiment was carried out with a simple non-adjacent pattern. In this experiment (Experiment 4), the test phase was slightly modified to accommodate a categorization task. In particular, after a brief training session, participants were asked to categorize each stimulus according to the grammar they learned while their brain waves were being measured. Due to the nature of the oddball design and the categorization task, I expected to find a difference between grammatical and ungrammatical words in the P3 rare-minus-frequent difference waveform. P3 is a positive-going wave peaking at around

300 ms and indexes conscious access to working memory updates (Polich, 2007). In addition to the P3, I expected to find a higher amplitude LPC to novel words that violated the learned phonotactic constraint than novel words that satisfied it. The results showed that participants learned the simple rule at the behavioral level (as measured with d-prime, a sensitivity measure to rule violations). This rule learning was also reflected in the brain response to violations of the rule, indexed by the P3 rare-minus-frequent difference waveform and the LPC. I conclude that brief sessions of laboratory learning of phonotactic rules result in a neural commitment which is reflected in the different brain responses.

The next experiment (Experiment 5) was conducted to investigate whether different types of exposure (implicit vs. explicit) lead to different neural commitments. Following Experiment 4, the aim was to examine the neurophysiological correlates of explicit learning of a non-adjacent phonotactic pattern. Contrary to Experiment 4 where participants merely repeated grammatical exemplars without any explicit instruction, in Experiment 5 the participants had the rule explained to them before the training session. While recording EEG, participants were presented with words that were either well-formed or ill-formed according to the rule. Participants in Experiment 5 performed behaviorally much better than participants in Experiment 4, reflected in both sensitivity index and accuracy. However, explicit learners did not show a significant ERP response modulated by well-formedness. The results show that explicit learning engages processing mechanisms that lead to prediction models at the behavioral level – but at the neural level, it is a different mechanism than the implicit group. This suggests that implicit lab-learning experiments tap into the kind of unconscious, automatic learning that is characteristic of natural language acquisition.

The final experiment was conducted to examine whether the degree of locality (the computational complexity of a local pattern) in phonotactic learning affects encoding at the neural level. In this experiment, our aim was to examine the neurophysiological correlates of a more complex pattern compared to the one in Experiment 4. The aim was to see whether the brain response in Experiment 4 was indexing a local pattern violation or a non-local pattern violation. Participants were exposed to a non-local phonotactic pattern at the syllable level and tested in an artificial grammar learning paradigm. While EEG was recorded, they were presented with words that were either well-formed or ill-formed according to the rule. I found no sensitivity to the ungrammatical stimuli, indicating that learning had not occurred. Moreover, there was no significant ERP response other than the obligatory responses such as auditory evoked potential and readiness potential, and no response modulated by well-formedness (neither P3 nor LPC). These results show that the degree of locality in the pattern modulates both learning and brain response. Brief exposure to the nonlocal pattern does not lead to a measurable change in the brain’s predictive systems. I conclude that exposure to computationally different patterns results in both different knowledge states and different underlying neural commitments.

The specific patterns that were tested in this dissertation are well-motivated from the perspective of formal language theory and theoretical linguistics. The results from these experiments provide support that (i) human phonological learning mechanisms are restricted by computational constraints; (ii) linguistic constraints that are acquired in laboratory situations are directly channeled into incremental, real-time phonological predictive processing; (iii) the nature of the neural commitment and the underlying knowledge state depends on the type of learning and the computational complexity of the patterns.

1.1 Overview of the Dissertation

This dissertation is organized as follows: Chapter 2 discusses the computational background of the tested patterns (Subregular Hierarchy) and reviews the artificial grammar learning paradigm and psycholinguistic studies that used this paradigm for learning phonological patterns. The Neural Commitment Hypothesis and Event-related-potentials (ERP) in language research are also discussed in this section. Chapter 3 focuses on the series of experiments that were carried out to examine behaviorally the relative learnability distinctions between the two long-distance harmony patterns that differ typologically (attested vs. unattested) and computationally. The details of the methodology, results, and implications of the results of these experiments are presented in this section. Chapter 4 introduces the EEG part of one experimental condition in Experiment 3 which was designed to establish the neurophysiological signatures of phonological learning via MMN. In this section, the methodology, the stimuli, and the implications of the results are discussed in detail. Chapter 5 introduces the attempt to index phonological learning via LPC response with a different design, namely with a different testing phase. The aim of this experiment was to observe an LPC to a phonotactic rule violation in the absence of semantic information, using a sibilant harmony rule. Chapter 6 describes an experiment testing a simple non-adjacent pattern, which was designed to explore the nature of neural commitment, shown to index phonological pattern violation. Chapter 7 describes an experiment testing the neurophysiology of explicit learning. The background of the distinction between implicit vs explicit learning and the details of the methodology are presented in this section. Chapter 8 introduces an examination of the neurophysiological correlates of a computationally more complex pattern. A general discussion and conclusion are provided in Chapter 9.

Chapter 2

BACKGROUND

In this chapter, I will broadly introduce the computational complexity assumptions behind the phonological patterns that are tested in this dissertation. First, I will introduce the Subregular Complexity Hypothesis and review a behavioral study (Lai, 2012, 2015) which will give a background about the details of the formal languages that I test the learnability assumptions of (section 2.1). I next review the artificial grammar learning paradigm and behavioral studies related to phonology/phonotactics (section 2.2). Then, I will talk about the basic principles and application of EEG/ERPs in language research and introduce the main ERP components reported in this dissertation (section 2.3). Lastly, section 2.4 summarizes the chapter.

2.1 Subregular Hierarchy and Complexity Hypothesis

The Chomsky Hierarchy (Chomsky, 1956) divides all logically possible patterns into nested regions of complexity (see Figure 2.1 below). Each of these regions has multiple mathematical definitions that enable any machine or algorithm to generate the strings comprising the pattern (Harrison, 1978; Hopcroft, Motwani, & Ullman, 2006). Also, each region specifically distinguishes abstract, structural properties of grammars: e.g., a machine with finitely many internal states can recognize patterns belonging to only the regular region. Phonological patterns belong to the regular region in this hierarchy (Johnson, 1972; Kaplan & Kay, 1994). Heinz (2010) further shows that phonotactic patterns in natural languages inhabit proper subsets within the regular region. These patterns are Strictly-Local (SL), Strictly-Piecewise (SP), and Non-Counting (NC or Locally Testable with Order) patterns (Heinz, 2010; Heinz & Rogers, 2013; McNaughton & Papert, 1971; Rogers et al., 2010; Rogers & Pullum, 2011).

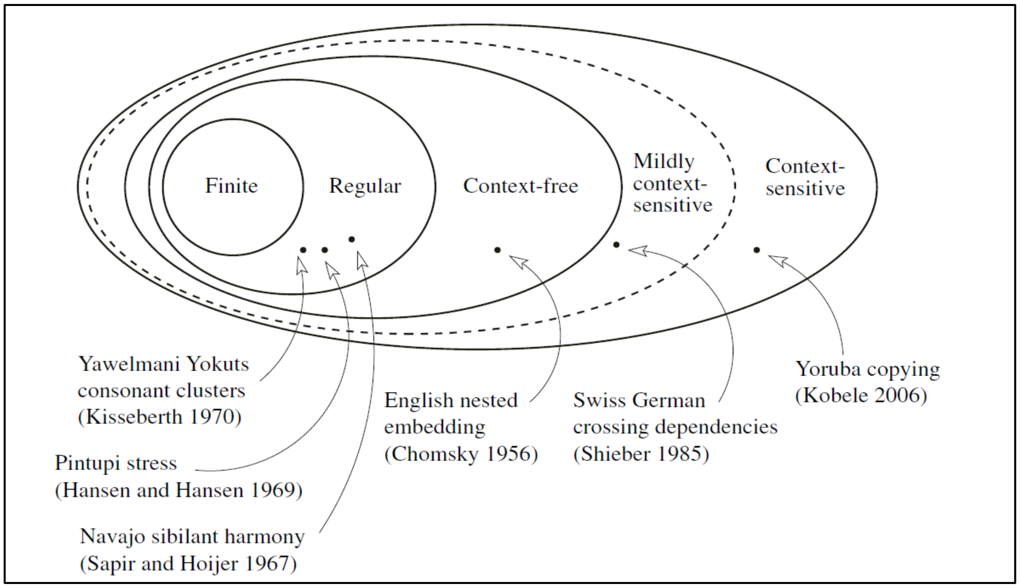


Figure 2.1: The Chomsky Hierarchy. Various features of natural language occupy different regions of the hierarchy. Figure reproduced from Figure 1 in Heinz (2010: p. 634) with permission.

A strictly *k*-Local (SL*k*) pattern is one where the well-formedness of a string is determined by whether or not its contiguous substrings of length *k* are well-formed. That is, SL languages only make distinctions on the basis of contiguous substrings up to some length *k* (called *k*-factors). The class of strictly *k*-local languages is known to represent the phonological patterns of spreading and correspondence restrictions in natural languages. A strictly *k*-Piecewise

(SP*k*) pattern, on the other hand, is one where well-formedness of a string is determined by its subsequences of length *k*: that is, if the set of subsequences in the string in question is a subset of the set of subsequences allowed by the grammar, the string is well-formed; otherwise, it is not. Thus, subsequences are not necessarily contiguous and the patterns they describe contain long-distance dependencies. The class of strictly *k*-piecewise languages is known to represent the phonological patterns of symmetric and asymmetric long-distance patterns like consonantal harmony. One example is a Sibilant Harmony rule that requires the agreement of features throughout the word, an attested rule in Chumash and Navajo native American languages (Applegate, 1972; Sapir & Hoijer, 1967) belongs to the SP region.

In the regular region, apart from SL and SP, there are also other regular patterns which are neither SL nor SP. These patterns can be collapsed under the Non-Counting patterns, which also goes by the name Locally Testable with Order. A pattern is Non-Counting if there is a number *n* such that for all strings u, v, w, if uv^nw occurs in L, then $uv^{n+1}w$ occurs in L as well (McNaughton & Papert, 1971). According to Heinz (2018), a simple rule named First-Last Assimilation that requires agreement between the first and last segment of a word specifically belongs to the Locally Testable (LT) class under the Non-Counting region (Heinz, 2018).

Figure 2.2 below presents a schematized representation of the subregular constraints and classifies attested and unattested phonological constraints as follows: Nasal Place Assimilation, an attested local dependency pattern; Sibilant Harmony (SH), an attested long-distance dependency pattern; and First-Last

Assimilation (FL), an unattested, non-SL, non-SP long-distance dependency pattern.

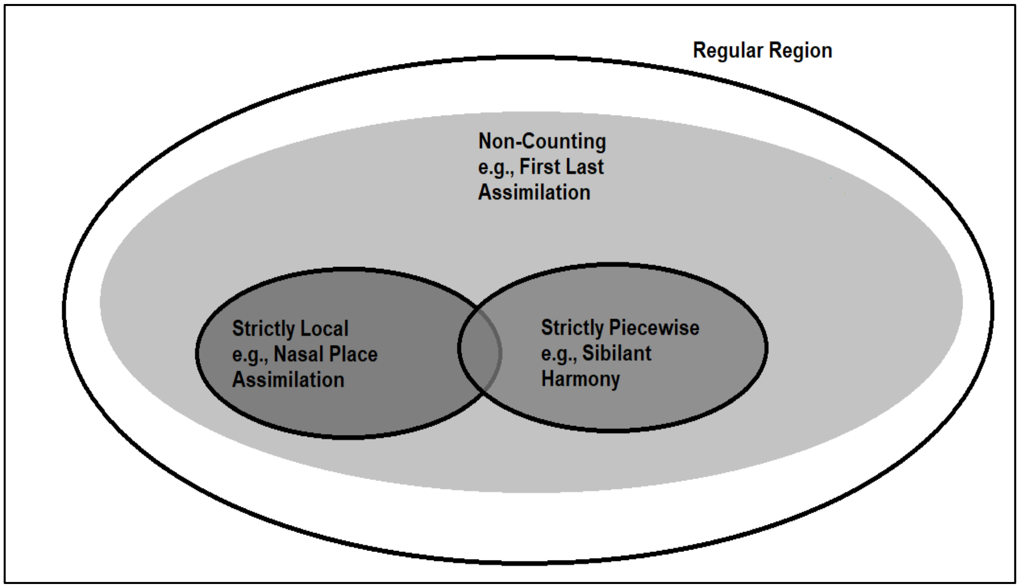


Figure 2.2: Subregular Boundaries. SL, SP, and Non-Counting are all in the regular region.

In contrast to the Non-Counting patterns, the SL and SP classes include almost all natural language phonotactic patterns (Heinz, 2010); that is, no language has a phonotactic pattern like FL. In this respect, Heinz's (2010) Subregular Hypothesis is supported by the typology of phonotactic patterns and predicts that humans' phonological pattern detectors can only learn phonotactic constraints that are SL or SP. If this is the case, then the absence of patterns such as FL from the natural languages can be explained; namely, the regularities present in patterns of FL cannot be extracted by humans' phonological learning mechanism. In other words, patterns with specific subregular computational properties are privileged with respect to learnability.

This learnability hypothesis was first tested by Lai (2012, 2015), a study which I partially replicate in a study reported in Chapter 3. Lai (2015) provided empirical evidence for the Subregular Hypothesis by using an artificial language learning paradigm to compare the learnability of two long-distance harmony patterns that differed computationally and typologically. The study compared the learnability of a Sibilant Harmony (SH) rule, an attested long-distance harmony pattern in the Navajo and Chumash languages, to a First Last Assimilation rule, an unattested long-distance harmony pattern. In Navajo, sibilants in a well-formed word have to agree in anteriority. Hypothetical words such as [sokosos] and [ʃokoʃoʃ] are grammatical because in both words each of the three sibilants has the same anteriority feature; whereas *[sokosoʃ] or *[ʃokosoʃ] are not grammatical because both words contain sibilants that disagree in terms of anteriority. On the other hand, First-Last Assimilation (FL) allows disharmonic intervening segments so long as the first and last sounds are harmonic. Therefore, [sokoʃos] and [ʃokosoʃ] are grammatical since in each word the first and the last sibilant agree in anteriority; whereas *[sokosoʃ] or *[ʃokoʃos] are not grammatical because the first and last sibilants in each word don’t agree in anteriority.

Lai (2015) employed an artificial grammar learning paradigm to test whether SH or FL can be learned by normal human subjects. The subjects participated in this study were sixty-six monolingual adult native speakers of American English. Three experimental groups were formed (SH, FL, and control); each of them had two phases: a training phase and a testing phase. The SH group was trained by listening to words that conformed to the Sibilant Harmony grammar, and the FL group was trained by listening to words that conformed to the First Last Assimilation grammar. The control group didn’t receive any kind of training. Participants in the SH and FL groups were instructed to listen and repeat each word after it was presented. The training phase (15 min long) was followed by a testing phase (10 mins long) in which participants’ task was to judge whether the first word or the second word in a given pair were more likely to belong to the artificial language they had just heard during the training. Participants in the control condition were asked to judge whether they liked the first word or the second word of each pair better as a possible word. All participants were given the same test stimuli. Results of this study showed that SH was learned by the participants, providing strong evidence that the SH grammar that is in the SP region was internalized. On the other hand, FL participants didn’t learn the FL grammar but also showed preference to SH (hypothetically they may have learned SH through FL). The reason for this unexpected preference was that anything that violates FL also violates SH, and anything that conforms to SH also conforms to FL (see Figure 2.3 below). And since the training stimuli in the FL condition included stimuli that also conformed to SH, participants may have developed an unintended bias to SH. In order to address this potential SH bias, a follow-up study was designed in which FL participants were trained with only FL stimuli, [sokoʃos] and [ʃokosoʃ]. Results of this follow up study showed that when participants were trained with intensive FL stimuli, they performed worse than control participants. In other words, after removing the ambiguous stimuli, FL group did not show any sign of learning, nor had any preference. The study then concluded that FL was harder to learn than SH because there was insufficient evidence to claim that FL was successfully learned. These results were consistent with the hypothesis that the phonological learner is restricted by the sub-regular constraints.

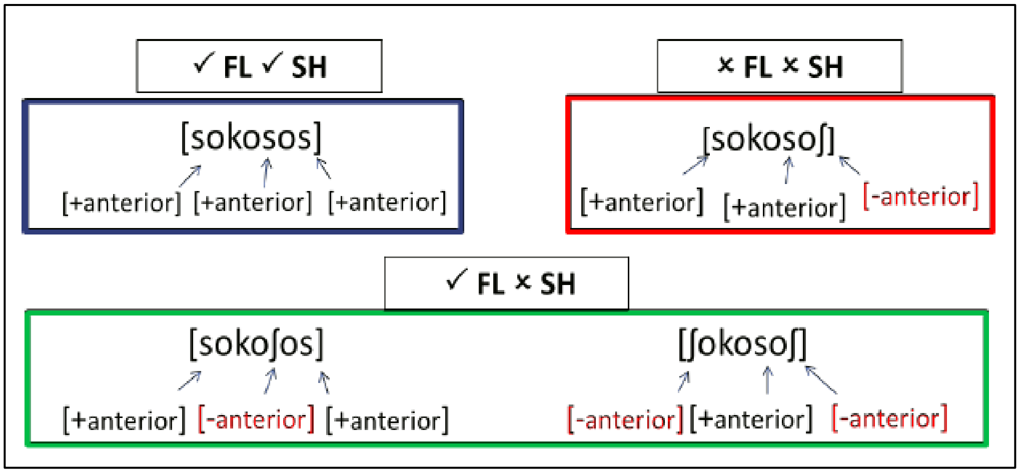


Figure 2.3: Comparison of SH and FL stimuli

2.2 Artificial Language Learning Paradigm

This dissertation employs the artificial language learning paradigm in order to test the availability of abstract levels of information during language learning. The paradigm itself dates back to Reber (1967). In a typical artificial language learning experiment, there are two phases: a training phase and the testing phase. In the training phase, participants are exposed to an artificial grammar constructed by the researcher. This artificial grammar is supposed to follow a pattern or a rule, and in the training phase, participants are expected to extract the pattern from the words or sentences of the grammar. The training phase’s aim is to replicate natural language acquisition as much as possible; therefore, no explicit feedback is given, and participants are not exposed to negative evidence – namely, the ungrammatical data. After the training phase, participants enter the testing phase in which they are being tested on whether they learned the pattern or the rule system of artificial grammar. Depending on the research question that is being asked by the experimenter, participants are also tested on whether or not they can generalize the pattern they learned to novel contexts that were not part of the training phase.

The artificial language learning paradigm offers a very good hypothesis testing space to linguists and cognitive psychologists who are interested in language learning. The number of studies using artificial language learning paradigm is growing (Finley, 2011, 2012; Finley & Badecker, 2009a; Hudson Kam & Newport, 2005; Kam & Newport, 2009; Moreton, 2008, 2012; Moreton & Pater, 2012a, 2012b; Pycha, Nowak, Shin, & Shosted, 2003; Wilson, 2003, 2006). In the following paragraphs, I will focus on the pros and cons of this paradigm.

One of the biggest advantages of this paradigm is that it enables the experimenter to control some factors that cannot be controlled in a natural setting such as the vocabulary or the sound inventory of a language. This enables the experimenter to test different languages that differ minimally (Finley, 2017). One other advantage of this paradigm is that it makes it possible to test patterns that are impossible to test in natural settings. For example, the learnability of unattested patterns cannot be tested by studying children throughout the process of normal language acquisition. On the other hand, this paradigm allows any language to be created and tested in a very controlled way which makes it possible for the experimenter to test a language pattern that is rare, extinct, or unattested. The learnability of patterns in a language like Haida, an endangered language that currently has 14 native speakers and is spoken in the Haida Gwaii archipelago off the coast of Canada and on Prince of Wales Island in Alaska; or like Latin, an extinct language; or like Na’vi, a constructed language spoken by the fictional Na’vi people in the world of James Cameron’s *Avatar*, can be tested in the artificial language learning paradigm.

However, this paradigm has also been criticized in the literature. See Moreton and Pater (2012a, b) for an overview of the paradigm and its criticisms. Firstly, the nature of learning in the training session is still not known. The paradigm can by no means control the biases the learner brings from their first language or the language experience the learner has. It is also still not clear whether artificial learning is like second language learning or first language acquisition. Ettlinger, Morgan-Short, Faretta-Stutenberg, & Wong (2016) argue that the learning that takes place in artificial settings is not like first language acquisition, but it is like second language learning. In order to overcome this drawback of the paradigm, it has been suggested to have a control condition that can factor out the biases of the learner (Finn & Kam, 2008; R. Reber & Perruchet, 2003).

One other disadvantage of the paradigm is that the artificial learning environment is not natural at all. In a natural learning environment, children are exposed to multiple aspects of the language such as prosodic cues, syntactic cues or semantic cues. There are also many social factors that cannot be included in an artificial setting like social cues and attentional cues (Frank, Tenenbaum, & Fernald, 2013). Making reference in language use is among one of the central language functions (Pinker, 2013; Siskind, 1996; L. Smith & Yu, 2008). In natural settings, children learn their first language slowly over a long period of time. In artificial settings, participants are exposed to a list of words repeatedly within a short period of time. And generally, these words are stripped of any information that could be useful for bootstrapping purposes; namely any existing resource that can be helpful for forming a category is removed (Lai, 2012).

Another criticism of the artificial language learning paradigm is the generalization problem. It is difficult to generalize the conclusions reached by the paradigm to real language learning. What if participants are not using their language learning mechanisms while they are in the training session because they are simply treating this problem as a language game, some kind of pattern learning game or a simply a math problem? If this is the case, since the experimenter cannot control what kind of learning mechanism the learners are using, it is difficult to generalize the conclusions drawn from any artificial language learning task. Although it is not a complete solution, according to Lai (2012) instructing the participants as though they will learn a foreign language before the training session can be a remedy.

All in all, the artificial language learning paradigm allows us to study a learning experience that is not possible to observe in natural settings; thus, the experiments in this study adopt this paradigm as a tool to explore the learnability of phonological patterns.

2.2.1 Phonological and Phonotactic Experiments

In the literature, there are many studies that employ the artificial language learning paradigm on the learnability of grammar (Marcus, Vijayan, Rao, & Vishton, 1999; Öttl, Jäger, & Kaup, 2015) and on phonology (e.g., Finley, 2017; Lai, 2015; Peperkamp, Le Calvez, Nadal, & Dupoux, 2006). It has been shown that after a very brief training session, both seven-month-old (Marcus et al., 1999) and sixteen-month-old infants (Chambers, Onishi, & Fisher, 2003), as well as adults are able to learn the grammar and the phonotactics of an artificial language (Onishi, Chambers, & Fisher, 2002).

Several other studies have shown that both adults and young children can easily extract regularities exhibited in an artificial language without explicit instruction (Aslin, Saffran, & Newport, 1998; Berent, Steriade, Lennertz, & Vaknin, 2007; Chambers et al., 2003; Gomez & Gerken, 1999; Gómez & Lakusta, 2004; Moreton, 2008; Peperkamp S., 2006; Plante, Gomez, & Gerken, 2002; Saffran, Aslin, & Newport, 1996; Tessier, 2006). Folia, Uddén, de Vries, Forkstam, & Petersson (2010) after reviewing the literature about the artificial grammar learning paradigm, conclude that there isn't any distinction between adults and children in terms of learning results.

There are also many artificial language learning studies that have tested the role of phonological or phonetic naturalness in learning (see Moreton & Pater, 2012a, 2012b for a review). It has been shown that both natural (Finley, 2012, 2017; Pycha et al., 2003; White & Sundara, 2014; Wilson, 2003, 2006) and unnatural patterns (Linzen & Gallagher, 2014; Peperkamp & Dupoux, 2007; Seidl & Buckley, 2005) can be learned after a training session.

Phonotactics is about permissible sound combinations in language. While many types of phonotactic patterns present dependencies between adjacent (local) segments (e.g. nasal spreading, voicing or place assimilation in consonant clusters), there are many phonotactic patterns that result from interactions between non-adjacent (non-local) segments with intervening elements (e.g. consonant or vowel harmony patterns). The learnability of adjacent patterns has been studied by Aslin et al. (1998), Chambers et al. (2003), Dell, Reed, Adams, & Meyer (2000) Goldrick (2004), and Onishi et al. (2002). In most of these studies, it has been observed that both infants and adults use phonotactic regularities (based on the transitional probabilities of syllables within words and between words) to segment words from a continuous stream of an artificial language. These studies conclude that adjacent patterns can be learned by employing statistical learning methods in artificial language learning settings.

Although the learnability of non-adjacent patterns poses different challenges to the learner due to their inherent complexity, it has been shown that they are also readily learned (Cleeremans, Schreiber, & Mcclelland, 1989; Finley, 2011, 2012, 2015; Finley & Badecker, 2008, 2009a, 2009b; Koo & Callahan, 2012; Moreton, 2008; Newport & Aslin, 2004; Onnis, Monaghan, Richmond, & Chater, 2005; Pycha et al., 2003; Wilson, 2003).

Pycha et al. (2003) compared the learnability of a vowel harmony rule to a vowel disharmony rule – the latter of which is not frequently found in human languages. In both the harmony and disharmony rules, the pattern that was being tested was a *C.V.C.V.C* where the *C.V.C* was the root and the *V.C* was the suffix.

The backness feature of the first vowel was used to predict the feature of the suffix vowel. The results showed that participants learned both the harmony and disharmony patterns, and there was no significant difference between the two. The study also tested whether participants could learn an arbitrary vowel dependency rule and showed that the learnability of the arbitrary rule was worse than the harmony and disharmony rules. Similarly, Wilson (2003) found that when participants were tested on assimilation and dissimilation processes compared to a random process, they were better at learning the former. However, no significant difference was observed between the assimilation and dissimilation processes. Both of these studies show that when participants are tested on unnatural patterns, they learn the natural patterns better.

In a statistical learning experiment, Newport & Aslin (2004) compared the ability to segment a word from a continuous speech stream using transitional probabilities. Two conditions were tested: (i) non-adjacent syllables; (ii) and non-adjacent segments. Results demonstrated that participants successfully segmented words when the dependency was between two non-adjacent segments but not syllables.

Finley (2011) reported that when participants were tested on the learnability of a sibilant harmony pattern in which long-distance dependencies with different distances were controlled by the number of intervening segments, their learning was locally biased, meaning that less long-distance patterns were preferred over the more long-distance patterns. Lai (2012) discusses this as evidence of the subregular complexity hypothesis in that the usage of an SL learner is prioritized over an SP learner. In a follow-up study, Finley (2012) further showed that when participants were trained on long-distance patterns with varying complexity (again depending on the number of intervening elements between the two segments), they were able to generalize beyond the training space and learn the long-distance dependencies in an unbounded way.

In another study, Finley (2015) tested the learnability of a long-distance vowel harmony pattern. The results showed that when the pattern included a transparent vowel that makes the dependency more complex, participants had a hard time learning the harmony pattern. Without the transparent vowel, participants learned the vowel harmony better, but extra training made it possible to learn the pattern with a transparent vowel. Koo & Callahan (2012) also reported learnability results from long-distance harmony patterns. In this study, the dependency between [s] and [l] and [l] and [m] were tested in trisyllabic words with the structure of *C.V.C.V.C.V.* It was reported that participants preferred novel legal words over novel illegal words. The results suggested that the dependency between the two sounds is learned by ignoring the actual distances between segments. Similarly, Hwangbo (2015) compared the learnability of two harmony patterns that differ typologically (attested vs. unattested) and computationally (simple vs. complex). The results showed that participants were better at learning the less complex pattern consistently even though the computationally simpler pattern was unattested.

These studies together with Lai (2012, 2015) demonstrated that the learnability of phonological patterns in artificial language learning paradigms depends on many factors including naturalness, typology, and computational complexity.

2.3 EEG/ERPs in Language Research

As the studies contained in this dissertation are neurophysiological in nature, below I will first present the basic principles and the application of EEG/ERPs in language research. Then, I will describe the main ERP components reported in the experiments in this dissertation on phonological rule learning. The overview of the literature about each ERP component will be presented in each chapter where EEG/ERPs are used as neural measures.

Electroencephalography (EEG) is an electrophysiological monitoring method that records electrical activity in the brain over time. This method uses several electrodes (the number of the electrodes depends on the system that is being used) placed on the scalp that measures the electrical activity produced by large populations of brain cells. One of the biggest advantages of using EEG on top of behavioral measures is that EEG is a measure of covert processes during the online processing of speech. Behavioral measures are collected at discrete points in time and behavioral data is end-point data that does not reveal what happens in the brain prior to the behavioral response. As explained by van Hell, Fernandez, Kootstra, Litcofsky, & Ting (2018: p. 136) "… while behavioral measures necessarily reflect the cumulative sum of all of the processing in response to a stimulus, EEG dynamically measures the millisecond-by-millisecond neural response to the stimulus." Therefore, different processing mechanisms during the comprehension of phonotactic pattern detection can be uncovered by the EEG method.

Event-related potentials (ERPs) are the measured brain responses that directly result from a specific sensory, motor, or cognitive event (Luck, 2014). ERPs are segmented from raw EEG data through an averaging process which removes random brain activity, leaving only the relevant waveform (Rugg & Coles, 1995). ERPs are time-locked to an external event such as the onset of a sound or behavioral response, and reflect the regularities in electrical brain activity (see Fabiani, Gratton, & Coles, 2000; Handy, 2005; Luck, 2014, and Luck & Kappenman, 2011, for an introduction to ERP methods and analyses). For example, when the EEG is time-locked to the onset of the behavioral response in an experimental context that requires a behavioral response after a categorization decision, the ERP signal is created by the voltage changes. And this ERP will reflect a millisecond-by-millisecond record of the brain activity before, during, and after the motor response, as it unfolds over time.

ERP components include a series of positive and negative voltage deflections reflecting the processing of the stimulus. These components are categorized according to their polarity, latency, amplitude and scalp distribution. When electrical waves are positive-going, the component is labeled a positive-going component (P) and when they are negative-going, the label of the component is a negative-going component (N). Latency reflects the time course of the ERP signal, and it is defined as the time from the onset of the stimulus to the peak of the component (when a component reaches its peak amplitude) within a time window. ERPs are named according to their polarity and peak latency in terms of amplitude (e.g., the P300 is a positive-going component that peaks at 300 ms, and the N400 is a negative-going component that peaks at 400 ms). Moreover, components also have names reflecting their paradigm or function such as the MMN – mismatch negativity (which is observed when there is a mismatch) or ERN – error-related negativity (which is observed when an error is made). Another property of the ERP components is the scalp distribution. For example, Left Anterior Negativity (LAN) is defined as a negative deflection with a left anterior scalp distribution.

2.3.1 Main ERP Components in This Dissertation

In this section, I will provide a general introduction to the main ERP components reported in this dissertation, namely the Mismatch Negativity (MMN) response, Late Positive Component (LPC), and the P300. I will also refer to the auditory evoked potential and readiness potential briefly. Further details on ERP profiles of phonological rule learning or phonotactic pattern processing (MMN, P300, LPC) will be provided in the context of individual studies.

The MMN, an auditory event-related potential, reflects a change detection response generated in the auditory cortex (Alho, 1995; Näätänen, Paavilainen, Rinne, & Alho, 2007). It is functionally a neural surprise response that can be elicited independently of attention (Näätänen, 1985; Näätänen & Michie, 1979; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001; Winkler, 2007). It is elicited by various kinds of abstract changes in auditory stimulation such as tone pattern violations (Tervaniemi, Maury, & Näätänen, 1994), grammar violations (Aaltonen et al., 2008), abstract rule violations (Näätänen et al., 2007), and local rule violations (Wacongne et al., 2011). Since it is independent of subject’s attention, it can be elicited even during sleep (Sallinen, Kaartinen, & Lyytinen, 1994), as well as in coma patients (Kane, Curry, Butler, & Cummins, 1993), and is sensitive to language-specific knowledge (Näätänen et al., 1997).

The experimental paradigm that elicits the MMN is called the oddball paradigm, which can be defined as follows: a deviant stimulus infrequently appearing among the repeated occurrences of a standard stimulus. The infrequent stimulus itself does not elicit MMN (Sussman, Chen, Sussman-Fort, & Dinces, 2014). The MMN response is the difference wave which is obtained by subtracting the auditory event-related potential (AEPs) to standard stimuli from that to deviant stimuli; and it is observed as a negative difference score waveform at the fronto-central and central scalp electrodes (Näätänen et al., 1997). The MMN response usually peaks at 150–250 ms from the onset of the deviant stimulus (Amenedo & Escera, 2000; Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989; Sams, Paavilainen, Alho, & Näätänen, 1985; Tiitinen, May, Reinikainen, & Näätänen, 1994). An essential prerequisite for elicitation of the MMN is that the central auditory system must establish a representation of the consecutive standard stimuli before the occurrence of a deviant stimulus. Thus, the auditory system is able to form a representation of the repetitive aspects of auditory stimulation, which is called the memory trace that can persist for a couple of seconds (Huotilainen et al., 1993; Paavilainen, Jaramillo, Näätänen, & Winkler, 1999; Paavilainen, Jiang, Lavikainen, & Näätänen, 1993; Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996; Winkler, Karmos, & Näätänen, 1996; Winkler, Lehtokoski, et al., 1999). In addition, the MMN reflects the discrimination between any auditory stimuli, specifically phonetic distinctions within a language at the sensory level (Näätänen, 2001); and more abstract phonemic contrasts (Cheour et al., 1998; Näätänen et al., 1997; Sharma & Dorman, 1999; Shestakova et al., 2002). Lastly, as the magnitude of the auditory change increases, the amplitude of the MMN increases and the latency shortens (Näätänen et al., 2007).

Another ERP component that is relevant to the data presented in this dissertation is the P300 – an ERP component that was first reported more than 50 years ago (Sutton, Braren, Zubin, & John, 1965). Since then it has been established as a robust response reflecting discrimination of a target stimulus from the standard stimuli across modalities (Duncan-Johnson & Donchin, 1977, 1982). The P300 depends on the role of the stimulus probability and task relevance and is elicited in the classical oddball paradigm where presentations of sequences of repetitive stimuli are infrequently interrupted by a deviant stimulus. It is a positive-going component that peaks with a latency of 250 to 500 ms and is observed strongly in the parietal region. The latency range can vary depending on many factors such as stimulus modality, task, and age (Polich, 2007). According to Kutas, McCarthy, & Donchin (1977) and Magliero, Bashore, Coles, & Donchin (1984), categorization speed (the time to detect and evaluate a target stimulus) is indexed by P300 latency. Johnson (1993) has reported that the P300 is strongly observed over midline electrodes (Fz, Cz, Pz), and the amplitude increases from frontal to parietal regions. Moreover, P300 amplitude is inversely proportional to the number of attentional resources: when the task difficulty increases, P300 amplitude decreases (Kramer, Wickens, & Donchin, 1985). Therefore, there is a direct link between attentional resource allocation and P300 amplitude.

The most common hypothesis about the underlying processes that give rise to P300 comes from the context updating theory (Donchin, 1981).

According to this theory, the P300 reflects a mental representation revision process. In an oddball framework, an attention-driven comparison process checks the representation of the previous events in working memory, and if a change is detected, the attentional processes update the stimulus representation. This context update elicits the P300. If no change is detected, the attentional process maintains the context of the current mental model (Polich, 2007). Evidence for attention-driven construction of the memory trace comes from the influence of time-induced limitations on P300. When the time between two oddballs increases, P300 amplitude increases; whereas, short intervals between two oddballs cause smaller P300 components (Fitzgerald & Picton, 1984). If the time interval between two oddballs is higher than 6-8 seconds, the probability effect is lost which is the evidence of the memory trace (Polich, 1990).

Luck (2014) notes that in the time course of the P300 wave, there are many distinct ERP components (see Polich, 2007, for a review of the P300 wave). There are two P3 subcomponents with different scalp distributions: P3a and P3b. While P3a originates from the stimulus-driven frontal attentional mechanism, P3b is more related to the temporal-parietal task-relevant activity (Courchesne, Hillyard, & Galambos, 1975; Dien, Spencer, & Donchin, 2004; Donchin, 1981; Van Petten & Luka, 2012). In this dissertation, discussion about the P3 or P300 components refers to the P3b subcomponent that reflects aspects of voluntary target detection (Sussman et al., 2014). In fact, my P3 measurements were taken from the rare-minus-frequent difference waves, following Luck et al. (2009).

The Late Positive Component (LPC) is a positive-going component that peaks around 600 ms after the onset of the stimulus and stays positive for around 400 ms (Friedman & Johnson, 2000; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992). The LPC's scalp distribution is observed as a positive deflection at the central-parietal and posterior scalp electrodes. According to Rugg & Curran (2007), the lateral parietal cortex is the source of the neural activity that results in the LPC. Due to its peak at 600 ms, it is usually known as the P600 and reported to index a range of violations at the sentence level (syntactic violations) like agreement, phrase structure, subcategorization, and long-distance dependencies (for a review see Gouvea, Phillips, Kazanina, & Poeppel, 2010, and Morgan-Short, Steinhauer, Sanz, & Ullman, 2012). Syntactically ambiguous sentences and garden path sentences which require reanalysis have been found to elicit the P600 (Osterhout & Holcomb, 1992; Osterhout, Holcomb, & Swinney, 1994). The functional interpretation of the P600 is the evaluation or reanalysis of an abstract structural relation at the sentence level (Friederici, 1995; Kaan, Harris, Gibson, & Holcomb, 2000; Tanner, Grey, & van Hell, 2017). It also reflects the processing of task-related events that are unexpected or improbable (Coulson, King, & Kutas, 1998; McCallum, Farmer, & Pocock, 1984). Since the P600 is associated with syntactic reanalysis, I will use the label LPC in this dissertation. The LPC has been linked to episodic memory and is described as indexing recognition effects (R. Johnson, 1995). Rugg (1995), in a subsequent memory paradigm (memory retrieval task), showed that studied (old) words showed larger LPC than for new words, and an effect known as the difference due to memory. It has been reported that the amplitude of the LPC increases when consciously remembering words that are later recalled (Rugg et al., 1998; M. E. Smith, 1993).

According to Patel, Gibson, Ratner, Besson, & Holcomb (1998), the LPC also indexes musical structure violations and reflects knowledge-based structural integration processes. In addition, Núñez-Peña & Honrubia-Serrano (2004) found that the LPC indexes an anomaly detection process in sequences that are governed by rules (specifically rule violations in arithmetic tasks). The interpretation of Núñez-Peña & Honrubia-Serrano (2004) about the LPC is the most relevant to phonotactic violations because phonotactic patterns are also rule-governed. McLaughlin et al. (2010) tested L2 learners' processing of phonological violations (Finnish vowel harmony) and found a late positivity component (LPC), a positive-going potential beginning about 600 ms after the onset of the grammatical violation. A similar late positivity has also been observed for phonological violations such as a long-distance restriction on the place of consonants in CCVC words in German by Domahs, Kehrein, Knaus, Wiese, & Schlesewsky (2009). Similarly, Moore-Cantwell, Pater, Staubs, Zobel, & Sanders (submitted) also observed an LPC to violations of lab-learned phonotactic rules. Using an artificial grammar learning design, this study showed that novel words violating a phonotactic constraint elicited a larger LPC compared to the novel words obeying the constraint. This study demonstrated that native English speakers who were trained in laboratory settings on a non-adjacent voicing agreement rule ([dugi] and [kuti]) showed an LPC in response to violations of this rule (*[todu] and [*kigo]).

In the literature, there is a debate as to whether the language-related P600 and the P300 oddball effect are distinct components (Coulson et al., 1998; Gunter, Stowe, & Mulder, 1997). It has been claimed that the LPC (or the P600-like late positive ERPs) is a member of the P300 family (Dröge, Fleischer, Schlesewsky, & Bornkessel-Schlesewsky, 2016). Coulson et al. (1998) claimed that salience and task relevance influences the amplitude and latency of both P300 and P600 – thus they are not independent of each other. In another study, it was claimed that the P600 is more related to the timing of the response than to the stimulus presentation, just like the P300 (Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014). On the contrary, Osterhout (1999) showed that the P600 and P300 are distinct components in that they occur in different conditions with different timing and distribution. Frisch, Kotz, von Cramon, & Friederici (2003) addressed this question by testing aphasic patients with lesions including the basal ganglia and with temporal-parietal lesions. The results showed that both groups displayed a clear P300 but only the group with temporal-parietal lesions showed a P600. These results indicate that P300 and P600 are different components that have different neural sources.

In addition to the main ERP components (MMN, P3, LPC), this dissertation also refers to some subclasses of ERPs. The Auditory evoked potential (AEP) is a brain response that is time-locked to the onset of the sound and is used to follow the signal generated by the sound. It is an obligatory response originating from the brain responding to the onset of a sound, such as speech sounds or tones (Musiek & Baran, 2007). It is a pattern of voltage fluctuations lasting about 300 ms arising from the onset of the word, comprised of the N1 (a negative-going potential at 100 ms), P2 (a positive-going potential at 200 ms), and P3 (a positive-going potential at 300 ms). Another relevant ERP in this dissertation is the Readiness Potential (RP or BP (from German *Bereitschaftspotential*)). The Readiness Potential is a measure of the activity in the motor cortex of the brain indexing voluntary motor movement (see Jahanshahi & Hallett, 2003, for a complete review of the BP). Elicitation of RP will be used as a sanity-check for the experimental design where no ERP is observed.

What is the advantage of using ERPs to study phonological processing? Behavioral measures (as discussed above) reflect the end state of phonological processing – for example, the moment the phonological parser is done with the word. On the contrary, ERPs provide a millisecond-by-millisecond record of signal changes in the brain and provide information about the neural activation during phonological processing. Measuring the involuntary brain response to words that contain phonotactic rule violations, where the brain response precedes the behavioral decision, will result in better and more consistent data about what has been learned and what has not. In addition, ERPs provide new knowledge about whether linguistic constraints that are activated in laboratory situations are directly "channeled" into incremental, real-time phonological predictive processing.

Brain response data will, in fact, provide stronger and more direct evidence for learning than behavioral data alone. Language acquisition, whether in the laboratory or as part of human development, is essentially "implicit" learning, i.e. learning that does not come with conscious awareness of the rules.

When participants make conscious decisions about word forms, those decisions can be contaminated by bias, errors, and conscious reflections. However, if a rule has been acquired, this means that the phonological parser must incorporate it – how else could participants analyze a word and make decisions about grammaticality? This suggests that grammatical knowledge can and should be measured at the level of automatic processing mechanisms that implement such knowledge. If the studies discussed in this dissertation demonstrate that subjects behaviorally are able to identify rule violations, but no brain response is observed during auditory processing of violating forms (other than obligatory responses like AEP or RP), then this suggests that the knowledge is “explicit” rather than “implicit”, which would raise new questions about whether artificial language learning experiments reflect innate and unconscious constraints on linguistic forms. If attentive brain responses such as MMN or P3 to violations are observed, that would provide evidence that the linguistic constraints are “channeled” into phonological parsing mechanisms when participants tacitly formulate target phonotactic rules – which indeed would be expected if the experiments are tapping into innate human language acquisition mechanisms. This idea will further be developed in more detail below.

2.4 Chapter Summary

To summarize this chapter, I introduced the Subregular Hierarchy and the learnability hypothesis (Subregular Hypothesis) – certain patterns are privileged with respect to learnability, originating from the typological and computational assumptions regarding phonological patterns. The learnability assumptions are not derived from the different regions themselves but from the typological and computational aspects of those regions. I reviewed Lai (2012, 2015) which provided empirical evidence for the Subregular Hypothesis by using the artificial language learning paradigm. Next, I reviewed the pros and cons of the artificial language learning paradigm and behavioral studies that used this paradigm to study the learnability of phonological/phonotactic patterns. I then presented the basic principles and the application of EEG/ERPs in language research by explaining the main ERP components reported in the experiments in this dissertation on phonological rule learning. Lastly, I tried to answer the question of what the advantages of using ERPs to study phonological processing are.

Chapter 3

EXPERIMENT 1: BEHAVIORAL TESTING OF SUBREGULAR HYPOTHESIS

This chapter presents the methodology, the results, and discussion of a behavioral experiment (Experiment 1) that was conducted to replicate Lai (2012, 2015). The aim is to test the learnability assumptions of the Subregular Complexity Hypothesis, namely whether the learnability of specific phonotactic patterns is restricted to particular subregular regions of the complexity hierarchy. Lai (2012, 2015)'s findings will be re-assessed experimentally by a behavioral experiment using the same training paradigm but a different testing paradigm to see whether the effects are robust across paradigms. The learning performance of human participants on attested vs unattested long-distance patterns is compared. I am particularly interested in understanding whether unattested long-distance subregular patterns, illustrated by the First-Last Assimilation rule, can be learned. As will be explained in the following sections, these simple regular patterns encode certain types of long-distance dependencies which are not observed in any human language. In particular, I compare the performance of human participants on these patterns with their learning performance of attested long-distance patterns – e.g., the Sibilant Harmony rule, another instance of a long-distance subregular pattern.

This study consists of three experimental conditions. The first condition tests the learnability of the Sibilant Harmony (SH) rule, an attested long-distance harmony pattern that belongs to a learnable subregular region (Heinz, 2010). The second condition tests whether the unattested First-Last Assimilation (FL) rule will be learned, a long-distance harmony pattern that belongs to a subregular region hypothesized to be unlearnable (or difficult to learn) (Heinz, 2010). The third condition tests the learnability of FL under an "intensive" condition. The Intensive First-Last Assimilation (IFL) condition is similar to the FL condition except certain training items are omitted to emphasize others. Intensive FL specifically belongs to the Tier-Based Strictly Local (TSL) class which is a specific generalization of Strictly Local class. It is defined with the phonological tier. Similar to the SL class, TSL class can be defined with conjunctions of negative literals after non-tier elements are ignored. For example, in a hypothetical word like *sakasas*, when the vowels are ignored, the sibilant tier *s.s.s* holds a local relation which is a limited kind of long-distance behavior, as noted by Heinz (2018). See Heinz et al. (2011) for a more formal definition of TSL languages and proofs for several computational properties of the TSL class.

Since anything that violates FL also violates SH, and anything that conforms to SH also conforms to FL, and since the training stimuli in the FL condition included stimuli that also conform to SH, there is a possibility that participants may develop an unintended bias to SH. In order to address this potential SH bias, the IFL condition was formed following Lai (2015).

Each group was tested with an artificial grammar learning experiment that had two sessions: a training phase and a testing phase. During the training phase, participants listened to words that conformed to the attested or unattested patterns and were instructed to repeat each word orally once they heard it. The training phase replicated exactly the training phase of Lai (2015). In the testing phase, participants were presented with words in an oddball paradigm (a deviant stimulus infrequently appearing among repeated occurrences of a standard stimulus), and participants were asked to judge whether each word violated any pattern or rule of the artificial language they had just learned during training. Therefore, the task for the participant was to find the deviant stimuli by pressing a response box button to indicate his/her decision. Button presses made by participants to deviant stimuli (e.g., sibilant disharmonic words) were recorded, and sensitivity index (d') as a measurement of learnability (i.e., sensitivity to violations of the rule) and bias were calculated according to Signal Detection Theory (Macmillan & Creelman, 2004). Thus, I applied the oddball paradigm to an artificial grammar learning experiment. The experimental details are explained in section 3.1.

The results, presented in section 3.2, show that we replicated the earlier findings that an attested and computationally learnable pattern is inside the hypothesis space of humans' phonological pattern detectors. Additionally, it has been demonstrated that sensitivity levels for the unattested and predicted-to-be-unlearnable patterns revealed a psychological domain-general processing mechanism. These findings show that the Subregular Hypothesis's learnability assumptions are psychologically valid. Implications of these results are discussed in section 3.3. The chapter concludes by discussing a dual mechanism of processing: (i) the linguistic domain-specific mechanism which is constrained by the Subregular Hypothesis in that phonotactic patterns that reside inside of specific subregular regions are privileged to be learnable; and (ii) the psychological domain-general mechanism which is not constrained by the

Subregular Hypothesis, and helps participants figure out a computationally difficult pattern in a laboratory setting.

3.1 Method

3.1.1 Participants

A total of 72 (24 in three experimental conditions) University of Delaware students were recruited as participants and provided written consent in the experiment. Each participant received course credit for participation. 66 of the 72 participants were females and 6 were males (this imbalance arises from the fact that the population we convenience-sampled from, undergraduates in University of Delaware Linguistics Department, was overrepresented with women). All participants provided informed consent to participate. Six participants were left-handed, but we did not exclude left-handers, as most left-handed people have left-lateralized language function (Somers et al., 2015). The mean age was 22 (*SD = 4.32, range = 18 to 31*). None of the participants reported a history of hearing loss or speech/language impairments and all reported English as their first and only language.

3.1.2 Design

Sensitivity index (d') and bias (c) (see Section 3.5 for the details of how d' and c are computed) values are used as a dependent measure and compared across each experimental condition in a one-way analysis of variance (ANOVA) to see whether there was a main effect of sensitivity difference as a function of the grammar that was learned. This is a one-way ANOVA with three between-subject groups: SH, FL, and IFL. In addition to the main effect of sensitivity to the ungrammatical words, I was interested in the question of whether the pattern violating phoneme, the position of the violation or the type of pattern (only for FL groups) had an effect on sensitivity detection. Following this aim, within the SH group, the effect of the violating phoneme (/s/ or /ʃ/) versus the violation point (middle or end) was analyzed as a 2 x 2 within-subject comparison. Within the FL and IFL groups, again the effect of violating phoneme versus the type of pattern (type a or b) was analyzed as a 2 x 2 within-subject comparison. Table 3.1 below summarizes within-subject comparisons in each experimental condition.

Table 3.1: Within-subject comparisons in each experimental group

			Within-subject contrasts	
			IV: Phoneme	
Groups			[s]	[ʃ]
SH	**IV: Violation Point**	Middle	*SH2-s [ʃ.s.ʃ]	*SH2-ʃ [s.ʃ.s]
		End	*SH3-s [ʃ.ʃ.s]	*SH3-ʃ [s.s.ʃ]
FL and IFL	**IV: Violation Type**	Type a	*FL3a-s [ʃ.ʃ.s]	*FL3a-ʃ [s.s.ʃ]
		Type b	*FL3b-s [ʃ.s.s]	*FL3b-ʃ [s.ʃ.ʃ]

The abbreviations above in Table 3.1 were used to code each type of disharmonic word. For example, in *SH3-s, “*” signals the disharmonic stimuli, “2 or 3” signals the violation point (middle or end), and the phoneme “[s] or [ʃ]” signals the violating sibilant which can also be viewed from the sibilant tiers. In the FL conditions, the violation point is always at the end, therefore, we coded whether the violating sibilant is different from the middle sibilant (*FL3a) or not (*FL3b).

3.1.3 Stimuli

Since we are partially replicating Lai (2015), the same stimuli (recordings) were used. All of the training and test stimuli had threes syllables in the form of *CV.CV.CVC*. The consonants in the alphabet of the language were only [k,s,ʃ], and the vowels were [a,ɛ,ɔ,i,u]. Half of the training stimuli had a stop [k] as the second consonant and the other half had it as the third consonant. Therefore, the first and last consonants were always sibilants.

In the testing phase, disharmonic words for each condition had four different forms. For the SH condition, the disharmonic sibilant was either [s] or [ʃ], and the position of this sibilant was either the second or the third sibilant. For the FL conditions, the disharmonic sibilant again was either [s] or [ʃ], and this sibilant was either different than the second sibilant or the same. All of the words which had a disharmonic sibilant at the end ([s.s.ʃ] or [ʃ.ʃ.s]) had a stop (k) as the second consonant. Half of the words which had a disharmonic sibilant in the middle ([s.ʃ.s] or [ʃ.s.ʃ]) had a stop (k) as the second consonant and the other half had a stop (k) as the third consonant. Table 3.2 summarizes the types of training and test stimuli used.

Lai (2015) reports that the stimuli used in the experiments were natural speech. A graduate student with phonetic training who was a native speaker of Mandarin Chinese was recorded. Explicit training was given to the recorder to ensure that all stimuli were produced consistently but she was unaware of the experiment’s purpose. All vowels were pronounced as full vowels. Word stress (with the acoustic correlates of increased pitch and loudness) was placed on the penultimate syllable of all words, and the sibilant [ʃ] was pronounced with rounded lips. The mean duration of stimuli was 1013 ms; the longest stimuli were 1251 ms and the shortest was 884 ms.

Table 3.2: SH, FL, and IFL show the harmonic stimuli used in both training phase and test phase; and *SH, *FL, and *IFL shows the disharmonic stimuli.

	Experimental Conditions					
Sibilant Tier	SH	*SH	FL	*FL	IFL	*IFL
[s.s.s]	[s.s.s]	None	[s.s.s]	None	None	None
[ʃ.ʃ.ʃ]	[ʃ.ʃ.ʃ]	None	[ʃ.ʃ.ʃ]	None	None	None
[ʃ.s.ʃ]	None	[ʃ.s.ʃ] (*SH2-s)	[ʃ.s.ʃ]	None	[ʃ.s.ʃ]	None
[s.ʃ.s]	None	[s.ʃ.s] (*SH2-ʃ)	[s.ʃ.s]	None	[s.ʃ.s]	None
[ʃ.ʃ.s]	None	[ʃ.ʃ.s] (*SH3-s)	None	[ʃ.ʃ.s] (*FL3a-s)	None	[ʃ.ʃ.s] (*FL3a-s)
[s.s.ʃ]	None	[s.s.ʃ] (*SH3-ʃ)	None	[s.s.ʃ] (*FL3a-ʃ)	None	[s.s.ʃ] (*FL3a-ʃ)
[ʃ.s.s]	None	None	None	[ʃ.s.s] (*FL3b-s)	None	[ʃ.s.s] (*FL3b-s)
[s.ʃ.ʃ]	None	None	None	[s.ʃ.ʃ] (*FL3b-ʃ)	None	[s.ʃ.ʃ] (*FL3b-ʃ)

3.1.4 Apparatus and procedure

The experiment was programmed with E-Prime Professional software v. 2.0.10.356, running on a Dell desktop PC. The experiment was conducted inside a single-walled shielded sound booth in the Experimental Psycholinguistics Lab at the University of Delaware. The presentation of sound stimuli was executed with two free field speakers placed in front of the participants at a comfortable listening volume (loudspeakers placed at 45° angles approximately 1 m in front of them), and visual input was delivered through an LCD display placed on a table in front of the participants. The PST Serial Response box was used for recording behavioral responses.

The procedure for these experimental conditions consisted of two phases: a training phase and a testing phase. During the training phase, participants listened to grammatical words (according to the experimental condition they were in) and were instructed to repeat each word orally once they heard it. The training session contained 200 tokens (40 words x 5 repetitions), and the duration was approximately 15 minutes. Therefore, the training phase was an exact replication of the Lai (2015). The training was followed by a testing phase in which words were presented in an oddball paradigm, where a deviant stimulus infrequently appears among repeated occurrences of a standard stimulus. "Standards" are the words following the pattern (SH, FL or IFL), and "deviants" are the words violating the pattern (*SH, *FL or *IFL). Participants were presented with words in a continuous stream and were asked to "press the bad word button when you think the word you heard does not belong to the language you had just learned during training" (even if they could not articulate the rule). Therefore, the task for the participant was to find the deviant stimuli (ungrammatical words) by pressing a response box button to indicate his/her decision. No explicit feedback was given to participants during the test phase because this would provide additional learning feedback during testing.

Each participant was presented with a total of 528 trials in the testing phase: 432 words (72 words x 6 repetitions) as grammatical (standards) and 96 (48 words (12 x 4) x 2 repetitions) words as ungrammatical (deviants) words. The test phase was divided into two blocks, each of which was the same in terms of the standards and deviants. The stimuli were delivered continuously, with a random number (between 3 and 7) of standards between each deviant. The 264 trials in each block consisted of 48 deviants (18%) and 216 standards (82%). The total duration for both training and testing was about 50 minutes.

3.1.5 Data recording and analysis

In Signal Detection Theory, the subject's response can be classified in four different ways: hits, false alarms, misses, and correct rejections. In the current experiment, the signal detection task can be described by the following table (Table 3.3) which is just an illustration of the possible four scenarios:

Table 3.3: Signal detection scenario in the current context

		Stimulus type	
	Signal	Ungrammatical (the "signal")	Grammatical ("no signal")
Subject's decision	"I noticed a signal"	Hit	False alarm
	"There is no signal"	Miss	Correct Rejection

The signal in this experiment is the ungrammatical stimuli, thus, for the purpose of computing the sensitivity index, only hits and false alarms are needed. In the test phase, button presses made by participants to deviant stimuli were recorded. When the signal (ungrammatical words) was present and the participant detected it and reported seeing it, it was counted as a hit. The proportion of hits was calculated as $P(H)= Nhits/Ndeviants$, with N being the number of times that the event was observed. When the signal was absent, but the participant still thought s/he heard something and reported it (when grammatical words were presented, and the participant reported it as ungrammatical), it was counted as a false alarm. The proportion of false alarms was calculated as $P(FA)= Nfalsealarms/Nstandards$.

The sensitivity index d' was then derived from the hit and false alarm rates according to signal detection theory (SDT) (Macmillan & Creelman, 2004). It was calculated as $d'=Z(P(H))-Z(P(FA))$, where P hits is the hit rate, P false alarms is the false alarm rate and Z is the z-score for that particular probability. To avoid infinite values in the z-scores, hit and false alarm rates were adjusted to $1–1/(2N)$ when they were actually 1 and to $1/(2N)$ when they were actually 0, where N is the number of trials on which the proportion is based (Macmillan & Creelman, 2004)[1].

The bias measure (c) which is participants' bias towards finding a signal was also derived from the hit and false alarm rates according to SDT. The bias

---

[1] 11 values in the SH group, 1 value in the FL and IFL groups were corrected in this fashion.

measure *c* was calculated as; *c= (Z(P(H))+Z(P(FA))/2*. The bias measure *c* reflects the balance between the false alarms and misses. Thus, when the false alarms and miss rates are equal, *c* is equal to zero. This illustrates the advantage of using SDT: Participants may be biased towards thinking that most words are grammatical (which is natural in normal language acquisition); using SDT allows us to factor out this bias from the accuracy. In the present experimental context, when the false alarm rate is higher than the misses, c becomes positive; when the misses are higher than the false alarm rates, c becomes negative.

When participants cannot discriminate at all, hits would be equal to false alarms which gives *d'* = 0. Results higher than 0 show that there is sensitivity to the signal. Thus, in the context of our study positive *d'* means participants are sensitive to the rule violations (in the sense that there is at least some sensitivity to ill-formed words). Likewise, c shows the bias towards reporting the signal or not. Negative bias along with the presence of positive *d'* values indicates that there is bias towards no signal which means, in the context of this study, the rule is learned. This is how learning is defined using SDT within our experimental context.

d' and c values are used as dependent measures and compared across each experimental condition in a one-way analysis of variance (ANOVA) to see whether there was a sensitivity difference as a function of the grammar that was learned. And within each experimental condition, *d'* and c values were compared across four deviant patterns in a 2 x 2 repeated-measures analysis of variance (ANOVA) with two factors: phoneme (two levels: [s] and [ʃ]) and either violation point (where the violating phoneme occurs in a word, two levels:

middle and end) (for the SH condition) or violation type (which type the stimuli is, two levels: type a and b) (for the FL and IFL conditions) (see Table 3.1 above). This allows us to test the null hypothesis that there should be no difference between the two phonemes, or there should be no difference between the two violation points (or types). All significant ANOVA effects are reported with the partial η^2 effect size measure.

3.2 Results

3.2.1 The main effect of grammar

When participants cannot discriminate at all, *P(H)= P(FA)* and *d' = 0*. Inability to discriminate means having the same rate of saying "yes" when grammatical words are presented as when ungrammatical ones are offered. As long as *P(H) ≥ P(FA), d'* must be greater than or equal to *0* (Macmillan & Creelman, 2004). In order to see whether any learning occurred, one-sample t-tests against zero (which would indicate that nothing was learned) was conducted (see Bendixen, Prinz, Horváth, Trujillo-Barreto, & Schröger, 2008, for the same method). Descriptive statistics for the three experimental groups are presented in Table 3.4 below.

d' results for the SH condition showed that deviants were detected with a mean sensitivity of *1.283 (SD=1.20)*, which is significantly different from zero sensitivity (*d' = 0*), t(23)=5.22, p<.001, d=1.066, 1-β=0.999. As for the FL conditions, deviants were detected with a mean sensitivity of *0.216 (SD=0.26)* in FL (t(23)=3.98, p<.001, d=0.813, 1-β=0.986) and *0.242 (SD=0.22)* in IFL (t(23)=5.33, p<.001, d=1.089, 1-β=0.999) both of which are also significantly different from zero sensitivity. In addition, the median scores of the groups also show that each group's median score is above zero, thus, each group as a whole has shown detection ability.

Table 3.4: Descriptive statistics for the three experimental groups

Descriptives	*SH*		*FL*		*IFL*	
	d'	*c*	*d'*	*c*	*d'*	*c*
Mean	1.283	-0.435	0.216	-0.763	0.242	-0.654
SE	0.246	0.081	0.054	0.077	0.045	0.080
Median	0.839	-0.528	0.271	-0.732	0.251	-0.657
Min	-0.042	-0.868	-0.386	-1.599	-0.275	-1.321
Max	4.272	0.683	0.607	-0.085	0.678	0.093

Mean bias rates for the SH condition showed that deviants were biased with a mean rate of -0.435 (*SD=0.40*) which is significantly different from no bias (*c* = 0), t(23)=5.37, p<.001, d=1.097, 1-β=0.999. As for the FL conditions, deviants were biased with a mean rate of -0.763 (*SD=0.38*) in FL (t(23)=9.92, p<.001, d=2.027, 1-β=1.00, and -0.654 (*SD=0.39*) in IFL (t(23)=8.17, p<.001, d=1.668, 1-β=1.00 – both of which are also significantly different from no bias. Thus, *c* was always negative and significantly different than zero, which is expected as a result of the oddball paradigm. Since the frequency of the signal is rare, participants are automatically biased to report that they heard an ungrammatical word.

The results of the one-way ANOVA showed that there was a statistically significant difference between the experimental conditions for the sensitivity index (F(2,69)= 16.990, p<.001, η2=.330, 1-β=0.966). Pairwise comparisons revealed that sensitivity to the ungrammatical words was statistically significantly lower for the FL (t=5.11, p<.001, d=1.22) and IFL (t=4.98, p<.001, d=1.20) groups compared to the SH group. There was no statistically significant difference between the FL and IFL groups (t=0.12, p =.991). This shows that participants who were trained with the sibilant harmony rule had a significantly higher sensitivity to the words that violated the pattern. See the left panel of Figure 3.1 below for the visual comparison of *d'* values.

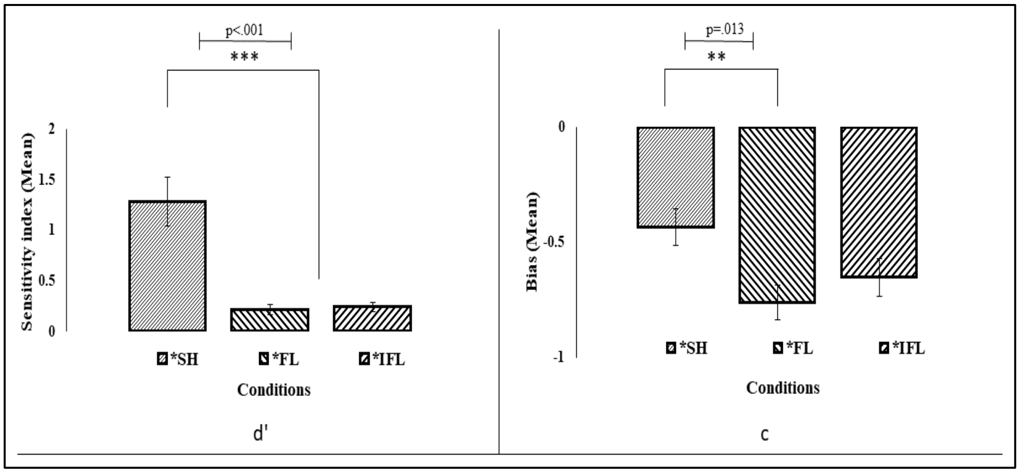


Figure 3.1: All Conditions: Group averages of sensitivity index rates of deviants (left panel), and bias rates of deviants (right panel). Error bars indicate standard error of the mean, stars indicate the significance of pairwise comparisons.

As for the bias rates[2], there was still a statistically significant difference between the conditions, as determined by one-way ANOVA (F(2,69)= 4.431, p=.015, η2=.114). Pairwise comparisons revealed that bias to the ungrammatical words was statistically significantly higher for the FL (t=2.92, p=.013, d=0.85) group compared to the SH group. However, there was no statistically significant difference between the FL and IFL groups (t=0.97, p = .598, d=0.28) or the IFL and SH groups (t=1.95, p=.132, d=0.55). See the right panel of Figure 3.1 for a visual comparison of c values.

3.2.2 Within-subject comparisons within the SH group

For the purpose of testing the null hypothesis that there should be no effect of phoneme difference (between the two phonemes), or violation point difference (between the two positions inside the word), we conducted 2x2 repeated measures analysis of variance using the sensitivity index as the dependent variable. The aim of these within-subject comparisons was to test whether the violations were harder to detect in the middle of the word; or whether the violating phoneme had any effect on detection. ANOVA results showed that the sensitivity index (d') was not significantly affected by the violation point or by the phoneme. See the left panel of Figure 3.2 for visual comparison. Also, there was no interaction between the violation point and phoneme (all p values >.05).

---

[2] We do not have a scientific hypothesis prediction for the bias rates because of the artifact of the oddball paradigm. Therefore, the only prediction for bias is the one that arise from the methods aspect—oddball paradigm.

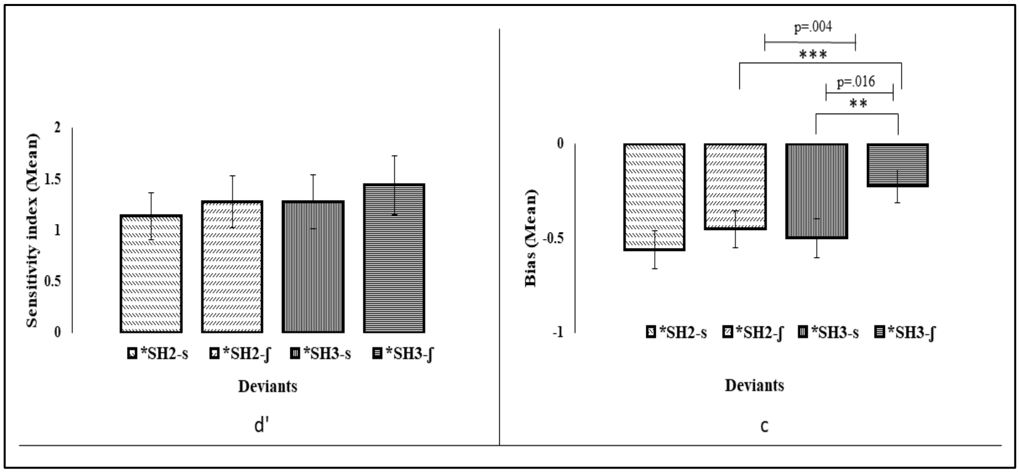


Figure 3.2: SH Condition: Group averages of sensitivity index rates (left panel) and bias rates (right panel) of ungrammatical words. Error bars indicate standard error of the mean, stars indicate the significance of pair-wise comparisons.

On the other hand, the bias rate was significantly affected by the violation point and by the phoneme such that the voiceless palatal-alveolar sibilant [ʃ], when used at the end of a word, attracted more bias (which was not predicted). The interaction between the violation point and phoneme was marginally significant. See the right panel of Figure 3.2 for a visual comparison. These results are summarized in Table 3.5 below. Pair-wise comparisons revealed a significant difference between *SH3-s vs. *SH3-ʃ (p=.016) and between *SH2-ʃ vs. *SH3-ʃ (p=.004). Other pair-wise comparisons were not significant (all p values > .05). This significant difference in bias could be attributed to the word edge saliency effect (Endress & Mehler, 2010). See below for further discussion.

Table 3.5: Repeated measures ANOVA statistics for the SH group along with F, p, η2 values, stars indicate the significance of within-subject comparisons.

	SH		
d' (Sensitivity)	F(1,23)	p	η^2
Violation Point	2.050	.166	.082
Phoneme	1.884	.183	.076
Violation Point × Phoneme	0.030	.865	.001
c (Bias)			
Violation Point	10.220	.004***	.308
Phoneme	4.439	.046*	.162
Violation Point × Phoneme	3.879	.061	.144

Reaction times (RT) were also analyzed to see whether there was a delay in the response time for specific phonemes or violations at specific points. Another 2x2 repeated measures analysis of variance analysis was conducted using the reaction time as the dependent variable. ANOVA results showed that the reaction time (d') was not significantly affected by the violation point (F(1,23)= 0.151, p=.702) or by the phoneme (F(1,23)= 2.718, p=0.113). These results show that participants while processing the word incrementally, were not significantly delayed by any phoneme (/s/ or /ʃ/) or the place of the violation (middle or end), see Figure 3.3 below.

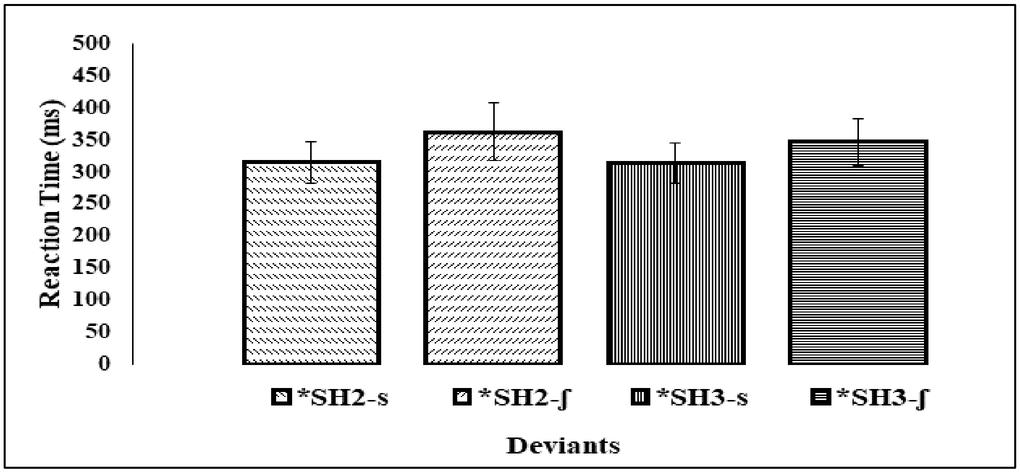


Figure 3.3: SH Condition: Reaction time differences across phonemes (s and ʃ) and violation points (middle and end).

3.2.3 Within-subject comparisons within the FL groups

The statistical test of the within-subject variables in the FL group shows that the sensitivity index (d') was not significantly affected by the violation type or the phoneme, and there was no interaction (all p values >.05). ANOVA results for the bias rates for the FL condition showed no significant effect of the violation type or the phoneme (both p values >.05). However, there was a significant interaction between the two. Such interaction also possibly arose due to the voiceless palatal-alveolar sibilant [ʃ] when used at the end of a word. See Figure 3.4 below. Moreover, pair-wise comparisons revealed a significant difference between *FL3a-ʃ vs. *FL3b-ʃ (p=.016). Other pair-wise comparisons were not significant (all p values > .05). Table 3.6 presents a summary of the statistical results.

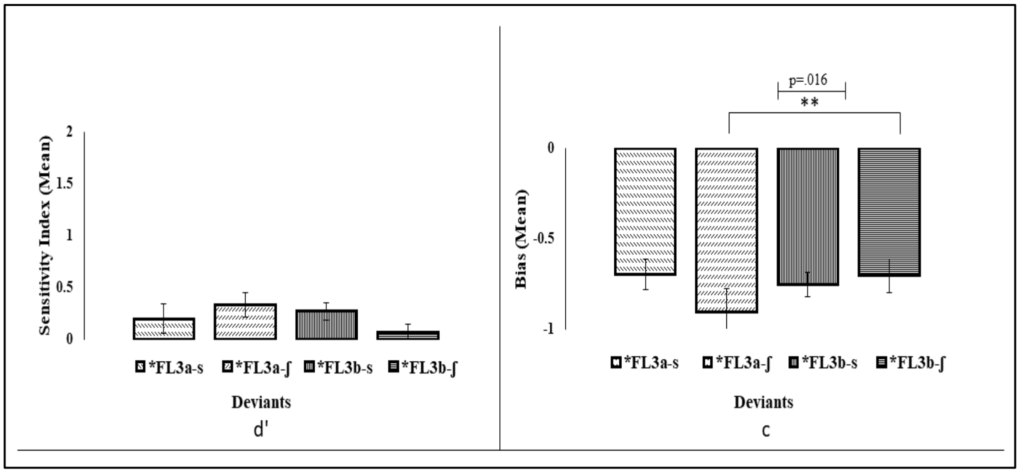


Figure 3.4: FL Condition: Group averages of sensitivity index rates (left panel) and bias rates (right panel) of ungrammatical words. Error bars indicate standard error of the mean, stars indicate the significance of pair-wise comparisons.

Table 3.6: Repeated measures ANOVA statistics for the FL groups along with F, p, η2 values, stars indicate the significance of within-subject comparisons.

	FL			IFL		
	F(1,23)	p	η^2	F(1,23)	p	η^2
d' (Sensitivity)						
Violation Type	0.780	.386	.033	3.589	.071	.135
Phoneme	0.085	.773	.004	0.002	.961	.000
Violation Type × Phoneme	3.210	.086	.122	1.705	.205	.069
c (Bias)						
Violation Type	2.286	.144	.090	0.014	.907	.001
Phoneme	0.715	.406	.030	3.007	.096	.116
Violation Type × Phoneme	6.556	$.017^{**}$	.222	0.006	.941	.000

In order to address the potential SH bias that was discussed in Lai (2015), an error analysis was conducted for the FL group. SH bias is that since a pattern that conforms to SH also conforms to FL, participants in the FL group, during their training of the FL patterns, might have developed a bias that helps them to learn the SH rule instead of the FL rule[3]. One possible way to analyze this issue is to look at the errors participants made during the test, to see whether there is a pattern that supports a possible SH bias. In the context of signal detection task, there are two errors: false alarms (signal was absent, but participant reported detecting it) and misses (signal was present, but the participant missed to report it). In terms misses, since all violations of FL pattern are at the same time violations of SH pattern, there is no way to differentiate the errors. However, the examination of the false alarms will definitely reveal whether there is an error pattern or not. When the violation is in the middle of the word, that poses a violation for the SH pattern but not for the FL pattern. That is, during the testing, when a word in the form of [s.ʃ.s] or [ʃ.s.ʃ] is presented, and FL learner should not press the button to report a violation. If the FL learner presses the button, that raises the question of whether the FL learner learned SH instead of FL due to the words that conform to both rules in the training.

To this end, false alarms were coded FLSH ([s.s.s] or [ʃ.ʃ.ʃ]) as a category to reflect the words that conform to both rules, and as FL_only ([s.ʃ.s] or [ʃ.s.ʃ]) to reflect the words that follow only the FL pattern. The mean

---

[3] Readers are referred to Lai (2015) that discusses the possible reasons of this bias.

probability of false alarm rates for these two categories was compared using a paired sample t-test. The results revealed that the mean probability rate for FLSH category was 0.218 (*SD=0.120*) and 0.228 (*SD=0.103*) for FL_only category. Results showed that there was no significant difference between the two categories, t(23)= 0.535, p=0.598, d=0.109. These results imply that the FL group did not have SH bias in that their error analysis shows that they do not have a significant preference for words that violate only the SH pattern.

For the IFL group, on the other hand, the *d'* or *c* was not significantly affected by the violation type or by the phoneme, and no interaction between the two was observed (all p values >.05) See Figure 3.5 below for visual comparisons.

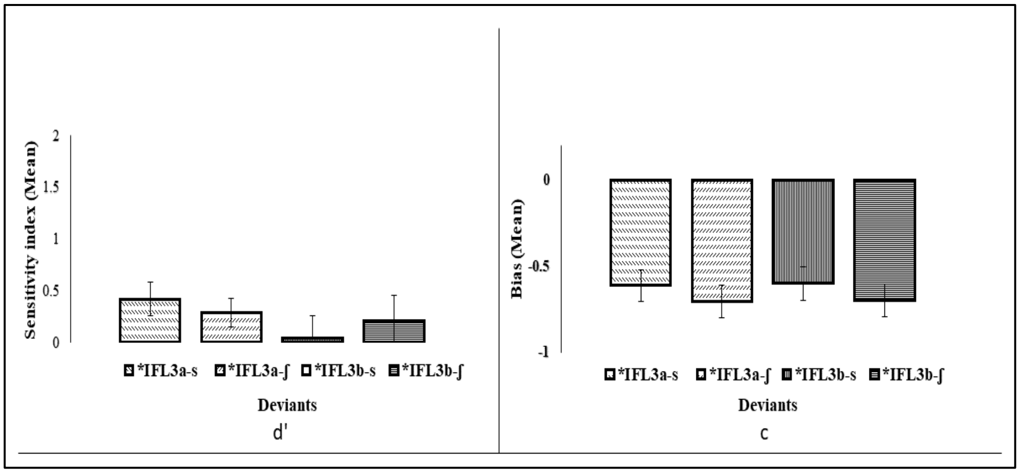


Figure 3.5: IFL Condition: Group averages of sensitivity index rates (left panel) and bias rates (right panel) of ungrammatical words. Error bars indicate standard error of the mean.

Table 3.6 above presents the repeated measure ANOVA statistics for within-subject comparisons of both FL groups. The only significant effect was the violation type and phoneme interaction in the IFL group. However, without a meaningful violation type or phoneme effect on its own, interpreting an interaction is not meaningful.

3.3 Discussion

By manipulating and violating the participants' auditory expectations at two distinct positions (or types) with two distinct phonemes, we aimed to obtain direct behavioral evidence which would support an active and predictive system underlying the brain's response to auditory stimuli.

The first objective of this study was to replicate Lai's (2015) domain-specific learning results in a different experimental paradigm (oddball design). The results above showed that in each experimental condition, participants showed sensitivity to the ungrammatical stimulus patterns, but at different levels of sensitivity. In the SH condition, all the deviants were detected with a mean sensitivity higher than zero and biased at a negative mean rate; thus, *sensitivity* for ungrammatical words was better than zero sensitivity which shows that participants learned the rule. The reason that we do not see a ceiling effect of d' like 4.65 (Macmillan & Creelman, 2004) is because of the computational complexity of the pattern – computationally simpler patterns would have resulted higher d'. Another point is that after the test session participants were informally asked what the pattern they learned would be, most of the subjects who showed good detection ability which is the implication of rule learning reported that they had no idea what the rule would be. This shows that they internalized the rule which is at the subconscious level.

Although disharmonic /ʃ/ at the end of a word was detected with a higher mean than the others, there was no effect of a phoneme or where the violation happens in a word on the sensitivity index. Reaction time results for the SH group did not reveal a processing delay when the violation is in the middle compared to the end. Together with the results of within-subject contrasts, this shows that during the incremental processing of words how the pattern is violated does not have an effect on pattern detection.

Furthermore, the results showed that participants were conservative and biased to report no signal for the disharmonic words, which is also an expected consequence arising from the probability of the signal (there are fewer signals than no-signals which are well known to lead to negative bias (Eschman, St. James, Schneider, & Zuccolotto, 2005; Hilgard, Weinberg, Hajcak Proudfit, & Bartholow, 2014)). Descriptive statistics for the bias rate suggest that disharmonic /ʃ/ at the end of a word made a significant difference. The reason for this could be the special status of word edges in phonology as discussed in Lai (2015): sounds at these positions become more perceptually salient. However, this is post-hoc speculation and has nothing to do with learnability or the Subregular Hypothesis. Pair-wise comparisons also showed that disharmonic /ʃ/ when occurring word finally was biased significantly differently. Although previous research showed that spectral peak is consistently measured higher for alveolar fricatives [s,z] than for palatal fricatives [ʃ,ʒ] (Soli, 1981), our results showed the reverse. Also, the perceptual salience difference between /s/ and /ʃ/

can be attributed to the frication itself, rather than to the surrounding vowel transitions (Harris, 1958; A. Martin & Peperkamp, 2011). Nevertheless, as we pointed out above since this study used the oddball paradigm in which the deviants were infrequent, a negative bias was not unexpected; thus, any conclusion based on the bias measure is speculative and not the focus of the current investigation.

The asymmetry is shown in the SH condition: positive *d'* and negative bias persists in the other two experimental conditions (FL and IFL) too. It seems that participants utilized the learning conditions to judge the grammaticality of the incoming stimuli. Although a kind of pattern knowledge was active at the behavioral level for the FL conditions – sensitivity rates significantly different from zero, the presence of an FL specific pattern learner cannot be implied. Nevertheless, participants in the SH condition were significantly *more* sensitive to the deviants, implying that unattested FL patterns are clearly more difficult to learn than the attested pattern.

The fact that participants in the unattested FL groups showed some sensitivity in the sense that their *d'* values are different than zero seems to contradict Lai (2015). However, the small sensitivity levels in the FL and IFL conditions do not falsify the Subregular Hypothesis's learnability claims. We argue that the small learning effect here is an artifact of the laboratory learning situation. We cannot prevent participants from employing general cognitive problem-solving procedures in a conscious way, and the small learning effect could simply be the result of active problem-solving strategies on behalf of the participants – i.e., domain-general learning such as relying on the saliency of word edges (Endress & Mehler, 2010). The observation that performance factors mask underlying knowledge is widely recognized in psychology (St. James, 2006). Thus, non-linguistic performance factors likely mask the innate constraints against FL rules during the experiment. A similar conclusion was reached in an fMRI study by (Musso et al., 2003), who trained participants to learn both linguistically attested rules in a language unknown to the participants, as well as linguistically unattested rules, violating principles of Universal Grammar. Although participants were able to behaviorally demonstrate in-laboratory learning of both rule types, only the UG-consistent syntactic rules activated Broca's area. We similarly predict that brain substrates relevant for linguistic phonotactics rules like SH would similarly show different activation patterns compared to brain regions responsible for general problem solving and FL-rule "learning."

The fact that participants showed significantly greater sensitivity to the SH pattern can be explained by the hypothesis of innate linguistic factors operative during learning, added on top of general psychological learning/problem-solving mechanisms. In this way, the Subregular Hypothesis can be thought of as an example of a domain-specific constraint on induction such that patterns which are attested in human languages exhibit constrained learning because they are channeled to language-specific learning modules.

3.4 Conclusion

In this chapter, I compared the relative learnability of two long-distance harmony patterns (Sibilant Harmony vs. First-Last Assimilation) that differ typologically (attested vs. unattested) and computationally (Strictly Piecewise vs. Non-Counting). I outlined an experimental framework for how abstract rules get translated into processing routines that generate real-time phonotactic predictions during auditory processing, and how this processing system is instrumental in pattern learning. I presented experimental results showing that adult learners behaviorally prefer some phonological patterns or distributions over others. These results substantiate the claim that the Subregular Hypothesis is active during real-time phonological parsing and to a significant extent constrains the learnability of specific phonotactic patterns. However, the fact that participants in the unattested FL groups showed some sensitivity shows that performance conditions can mask the underlying knowledge structure as described by the Subregular Hypothesis, in this case favoring making correct guesses about a non-linguistic pattern.

I needed to replicate the basic Lai (2015) findings in an oddball design in preparation for using this paradigm in an EEG experiment. EEG is a measure of covert processes during the online processing of speech. Behavioral measures are collected at discrete points in time and behavioral data is end-point data that does not reveal what happens in the brain prior to the behavioral response. Therefore, different processing mechanisms during the comprehension of phonotactic pattern detection can be uncovered by the EEG method. In the following chapters, I will present attempts to correlate phonotactic rule learning with an event-related potential (ERP).

Chapter 4

EXPERIMENT 2: CAN MMN INDEX PHONOTACTIC RULE VIOLATIONS?

This chapter presents the first attempt to correlate phonotactic rule learning with an event-related potential (ERP). In Experiment 1, the brain waves of the Sibilant Harmony (SH) group was recorded while participants were detecting words that violated the pattern. The behavioral results of that experiment were presented in Chapter 3, and the EEG results will be presented in this chapter[4]. It was observed that the SH pattern was learned at the behavioral level. Now, we will see whether the behavioral rule learning after brief exposure in laboratory settings caused neural commitment. Specifically, I will try to answer the question of whether the MMN indexes a phonotactic rule violation. Section 4.1 will present a relevant background and discuss the reasons for expecting MMN to phonotactic rule violations. The experimental details are presented in Section 4.2, and the results are given in Section 4.3. Section 4.4 discusses the results and the chapter concludes with Section 4.5 by discussing the limitations and directions for future research.

4.1 Background

Lai (2012, 2015) and our behavioral replication of this experiment showed at the behavioral level that the logically possible Sibilant Harmony rule

---

[4] Experiment 2 is not a new experiment, it just presents the EEG results of the SH group of Experiment 1. Brain waves of the FL groups were not recorded.

is learnable. When participants were exposed to words conforming to this pattern, they detected it. The reason for this was that SH lies inside the hypothesis spaces of humans’ phonological pattern detectors, contrary to unattested patterns like FL. In this chapter, I ask whether violations of SH stimuli trigger a brain response. Do the pattern detector restrictions extend into the perceptual system? At what level of processing is the subregular constraint active? If knowledge of phonological patterns is part of an internal grammar governing language perception and production, then do features of the grammar extend into the perceptual system?

Contrary to a local dependency, sibilant harmony is an example of a non-local dependency which means subsequences are not necessarily contiguous as can be seen in hypothetical words such as [sakesos] and [ʃakeʃoʃ]. Cognitive neuroscience evidence for local patterns has come from several studies (Conway, Pisoni, & Kronenberger, 2009; Furl et al., 2011; Näätänen, 2001; Näätänen et al., 2007) showing that the auditory system can determine relations between consecutive sounds given that the input is sequential. In this respect, MMN responses to violations of complex inter-tone relations can be taken as evidence of this capability (Bendixen et al., 2008; Saarinen, Paavilainen, Tervaniemi, & Näätänen, 1992). Also, the auditory system is shown to extract regularities (via MMN) from continuous sequences in which the tone relations keep dynamically changing (Bendixen et al., 2008; E. Sussman & Winkler, 2001).

As for the non-local dependencies, Bendixen, Schröger, Ritter, & Winkler (2012) have shown that the MMN is also elicited to violations of non-adjacent transitions by proving that the auditory system has the capacity to form links between non-adjacent elements in a tone series. Aaltonen et al. (2008) showed that detection of native-language rule violations is ‘cerebrally hardwired’, meaning that the linguistic difference between two languages (Finnish and Estonian) is reflected in the brain responses of native speakers of these languages. The study tested Finnish Vowel Harmony (FVH) violations by using the MMN paradigm, and the hypothesis was that Finnish brains should treat standard stimuli [tækæ] as normal but deviant stimuli [tækæ] as a violation of FVH, facilitating discrimination of the sounds. In contrast, for Estonian brains, the difference between [tækæ] and [tækæ] should not be particularly evident. The results showed that violations of a language-specific rule were automatically detected by Finns’ brains. This indicated that mere exposure to statistical regularities in one’s first language automatically results in the formation of language-specific neural networks (Aaltonen et al., 2008). Furthermore, MMN enhancement also caused by learning a second language in adulthood due to the formation of new phoneme categories (Tremblay, Kraus, & McGee, 1998; Tremblay, Kraus, Carrell, & McGee, 1997; Winkler, Kujala, et al., 1999; Winkler, Lehtokoski, et al., 1999). In another study, Cheour et al. (1998) found that sleeping newborns can learn to discriminate vowels after they were exposed to Finnish /y/–/ i/ contrast during sleep.

These results are parallel to those of Kuhl et al. (2008)’s native language neural commitment (NLNC) hypothesis which argues that language exposure produces neural commitment that affects future learning. That is, initial exposure to a language stimulus causes physical changes in neural tissue which in turn reflects properties of the language input. This means that if a rule like SH is learned after being a listener is exposed to it, this learning should lead to an instant "tuning" of the perceptual system that can be measured. A subsequent question in the present context is whether training can cause a physical change in the brain or how much language exposure is needed. According to Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen (2001), in order to establish sufficiently accurate memory traces that would be accessible to the early auditory processing mechanisms for the automatic change detection to occur, some training is required. Paavilainen et al. (1993) demonstrated that in one experimental session with active listening, a memory trace for the sound pattern is developed. In addition, statistical learning experiments in laboratory settings suggest that infants with just 2 min exposure to novel speech stimuli are able to learn (Maye, Werker, & Gerken, 2002; Saffran et al., 1996).

Taken all together, these findings demonstrate that MMN elicitation to violations of dependencies in harmony patterns would be a first step to take. However, it should be noted that most of the studies that show MMN elicitation in response to rule-governed sequences have used non-linguistic stimuli. The ones that used the linguistic stimuli (Aaltonen et al., 2008) used the same stimuli as standards to form the memory trace, and the same deviant stimuli to make the surprise effect. Also, to my knowledge no other studies used CV.CV.CVC type of multisyllabic linguistic stimuli. In sum, we recorded the brain waves of the Sibilant Harmony (SH) group in Experiment 1 while participants were detecting the words that violated the pattern. The aim was to test whether violations of the

SH pattern trigger a perceptual response (MMN) to novel stimuli after a training session.

4.2 Method

4.2.1 Participants

In this section, only the EEG related specifics about participants are presented. A total of 29 University of Delaware students were recruited as subjects: four subjects were excluded from analysis because they had experienced a different design[5], and one participant was excluded because of recording errors. The remaining 24 subjects had on average 92% good trials after artifact correction. 22 of the 24 subjects were females and 2 were males. One participant was left-handed. The mean age was 20 (*SD = 2.2, range = 18–31*).

4.2.2 Stimuli

The details about the stimuli were presented in Section 3.1.2 above. As a recap, only the SH stimuli were tested. Of the words conformed to Sibilant Harmony, half were in the form [s.s.s], and the other half were [ʃ.ʃ.ʃ]. Disharmonic words were in four different forms (each form had 12 words); (i)

---

[5] At first, I conducted a pilot study that had a different design. It included a sanity check condition (SCC). The stimuli in the SCC was like [sokosot] or [ʃokoʃot] to which we hypothesized to elicit very clear MMN responses because the SH pattern is violated by a stop sound. However, very early behavioral results showed that participants noticed all of the SCC violations easily (as we expected) but this caused the real violations to SH like *[sokosoʃ] being ignored. Therefore, in response to this unexpected SCC bias, we excluded the SCC from the experiment and the first four subjects from the analysis.

the disharmonic sibilant was [s] and it was in the middle (*SH2-s), (ii) the disharmonic sibilant was [ʃ] and it was in the middle (*SH2- ʃ), (iii) the disharmonic sibilant was [s] and it was at the end (*SH3-s), and (iv) the disharmonic sibilant was [ʃ] and it was at the end (*SH3- ʃ).

4.2.3 Design

This study tested the learnability of the Sibilant Harmony (SH) rule via MMN. The question of whether the pattern violating phoneme, or the position of the violation has an effect on the MMN amplitude was tested. The design included a 2 x 2 within-subject comparison: the effect of violating phoneme (/s/ or /ʃ/) versus the violation point (middle or end).

4.2.4 Apparatus and procedure

Apparatus and the overall procedure were presented in Section 3.1.4 above.

4.2.5 Data recording

Behavioral data recording and analysis were presented in Section 3.1.5 above. Electrophysiological recordings were made continuously from 128 carbon fiber core/silver-coated electrodes in an elastic electrode net (Geodesic Hydrocel 128). The acquisition and digitization of the data were performed with EGI Net Station software v.4.5. Electrode impedances were lowered to below 50 kΩ before data collection and the electroencephalogram (EEG) was continuously recorded with 24-bit digitization at 250 Hz. Each electrode was referenced to Cz and the analog recording passed through a 0.10 Hz first-order high-pass filter. After recording, the continuous EEG was segmented into epochs of 800 ms. Each epoch included a 200 ms pre-stimulus period before the stimulus onset (to be used for baseline correction), resulting in 600 ms of data for each single sound presentation. A 0.1 – 40 Hz bandpass filter was applied prior to latency measures. The next step was artifact detection in which bad channels, eye blinks, and eye movements were identified: if the difference between the maximum and minimum voltage exceeded 200 μV in a moving average of 80 ms, the channel was marked bad. If a channel is bad in over 20% of trials, it was considered bad in all trials; and trials containing more than 10 bad channels, eye blinks or eye movements were marked as bad. A spherical spline interpolation was used to replace bad channels. After that, each trial was baseline corrected using the mean voltage of the first 200 ms (Dien, 2010). The signals were re-referenced offline to the average voltage, which is the least biased reference method with high-density EEG.

4.3 Results

4.3.1 Behavioral Results

Behavioral results were already presented and discussed in Chapter 3 above. As a recap, participants learned the sibilant harmony rule with a *d'* of 1.283 (*SD=1.20*) which replicated the earlier findings showing that the SH pattern is inside the hypothesis space of humans' phonological pattern detectors, by testing violations of such patterns trigger behavioral responses to novel stimuli after a training session.

4.3.2 Electrophysiological Results

Electrophysiological results were analyzed in two different analyses.

4.3.2.1 Analysis I

I first tested the effect of phoneme on the corresponding voltage waves, since a possible phoneme difference, independently of the violation, would have an effect on subsequent comparisons. This is a test of the auditory evoked potentials (AEPs) for the first sound coming out of silence, and the beginning of the word, to see if there is a difference between /s/ and /ʃ/. Figure 4.1 below shows the auditory evoked potentials (AEPs), a pattern of voltage fluctuations lasting about 300 ms, arising from the onset of the word (left); and the spatial distribution of the mean voltage change of two phonemes from 100 ms before the onset of word to 300 ms after it (right).

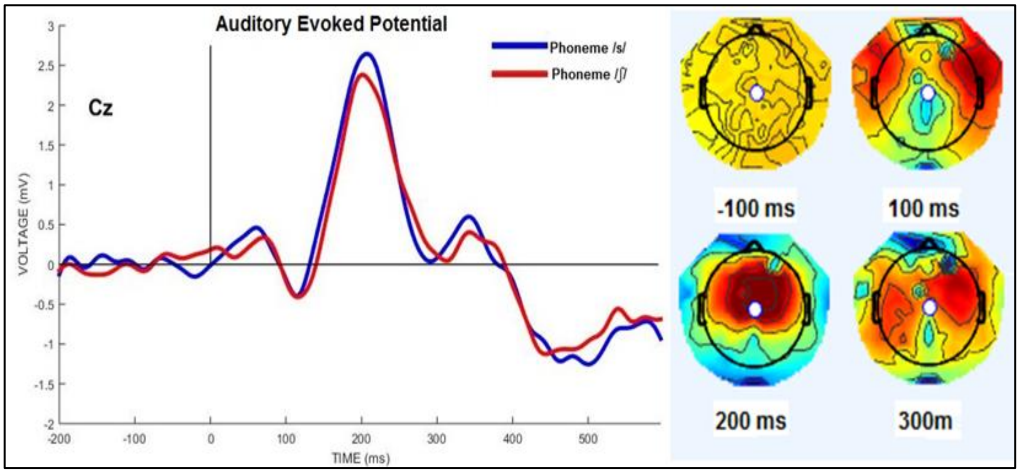


Figure 4.1: The specific neural activity arising from the response to the onset of the word out of silence (left panel), and spatial distribution of the mean voltage change (right panel).

It can be seen that there was a negative deflection at the central scalp electrode (Cz) 100 ms after the onset of the word which is reflecting the N1 component. This negativity was followed by a positive deflection around 200 ms (P2); and finally, around 300 ms the positivity decreased by giving place to negativity in the parietal area. Although according to the waveform, both /s/ and /ʃ/ patterned together, there was a small amplitude difference between the two. In order to see whether this difference was significant, a statistical analysis was conducted on Cz which was determined to show the biggest effect after visual inspection. An ANOVA was computed with the within-subjects factor phoneme (/s/ vs /ʃ/) on the mean voltages within a 200-600 ms time window after the onset of the word. The ANOVA result showed that there was no significant difference between the /s/ (M= 0.097 μV) and /ʃ/ (M=0.104 μV), F(1,23)= 0.30, $p< .05$. I conclude that the grammaticality effect of /s/ and /ʃ/ would be the same.

The next step was the analysis of any MMN-like response to the standard vs. deviant change in the stimuli for all four comparisons across the frontal-central region (see Section 4.2.3 above). Since there was no significant difference between (/s/ vs /ʃ/), the EEG comparisons were made over combined cells. For instance, *SH2-s and *SH2-ʃ were combined to form *SH2 as a single deviant. Figure 4.2 shows the ERPs obtained for standard and deviant words in the test sequence as well as the rare-minus-frequent difference waveforms, separately for when the violation is in the middle (SH2 vs. *SH2) and at the end (SH3 vs. *SH3).

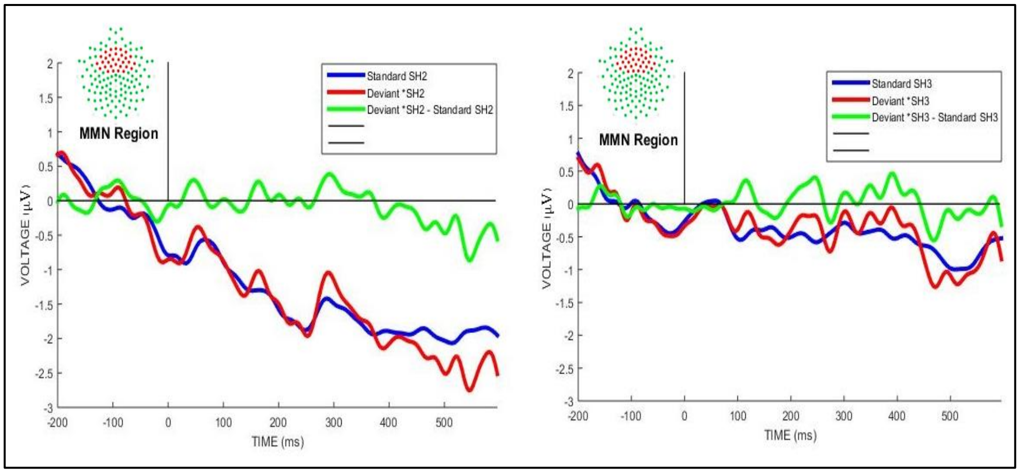


Figure 4.2: Time course of corresponding absolute voltage and difference waves in the grand average voltage for SH2 vs. *SH2 (left panel), and for SH3 vs. *SH3 (right panel).

Figure 4.3 below displays the ERPs obtained for standard and deviant words as well as the rare-minus-frequent difference waveforms for SH vs. *SH irrespective of whether the violation point was in the middle or at the end.

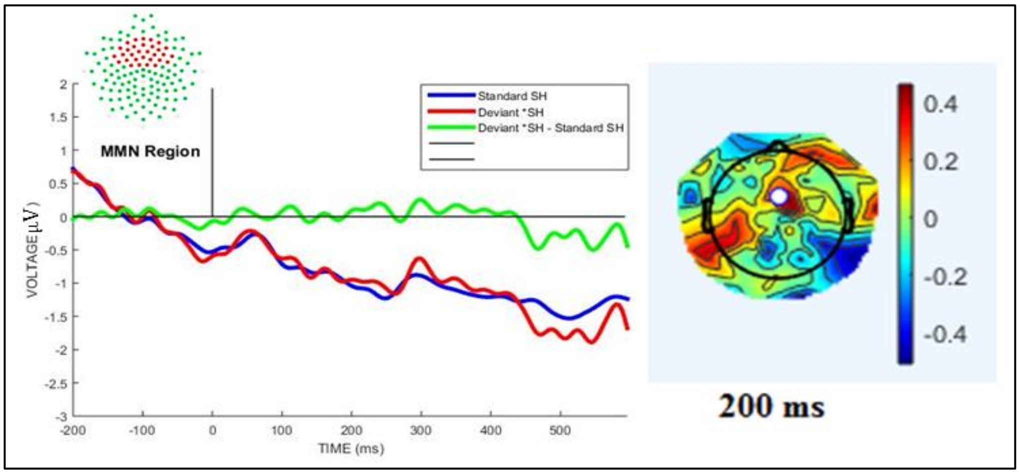


Figure 4.3: Time course of corresponding absolute voltage and difference waves in the grand average voltage for SH vs. *SH (left panel); and spatial distribution of difference wave at 200 ms (right panel).

Figure 4.3 also includes the topoplot which shows the mean difference wave at the horizontal line in the waveform plot at 200 ms. Both Figures show the waveforms time-locked to the violation point. In order to see whether there is a statistically significant difference between standard and deviant waves, frontal-central electrodes (see ROI in the figures) were selected, and ANOVAs were computed with three factors. The design was a 2*2*2 within-subject factorial repeated measures ANOVA in which each of the 3 factors has two levels; grammaticality (harmonic, disharmonic), phoneme (/s/, /ʃ/) and violation point (middle, end), and the dependent variable is the mean voltage (μV) change in the frontal-central region 100-300 ms after the onset of the violating sound. The results showed that the amplitude (μV) was not significantly affected by the grammaticality (SH vs. *SH), F(1,22)=.084, p< .05; and not significantly affected by the phoneme, F(1,22)=.023, p< .05. Although the violation point significantly affected the mean voltages, F(1,22)=15.878, p=.001, without a significant grammaticality effect, it is not meaningful. Also, there was no interaction effect between the variables, all p values >.05.

A correlation analysis was also conducted between behavioral sensitivity and MMN. Results showed that MMN amplitude was not correlated with the behavioral sensitivity, r= -0.313, p=0.136, see Figure 4.4 below.

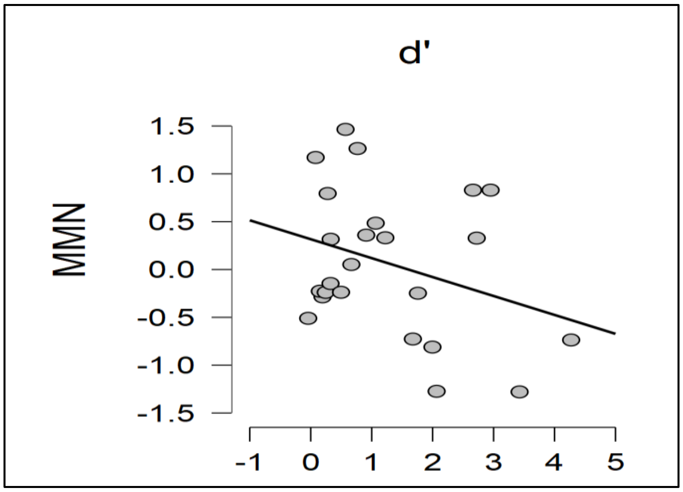


Figure 4.4: The correlation matrix between d' and MMN amplitude.

4.3.2.2 Discussion of Analysis I

The results in Figure 4.1 revealed a typical AEP with an N1–P2-P3 wave-form complex at Cz. It was important to observe that both /s/ and /ʃ/ at the onset of the word showed a robust AEP in the sense that the auditory system of participants responded to the onset of the word the same. Furthermore, the fact that there was no statistically significant amplitude difference between the phonemes was parallel to the behavioral results. This confirmed that the grammaticality effect of /s/ and /ʃ/ was the same across all words. If the statistical results had shown a significant phoneme effect to the onset of the word, [s.s.s] as standard and *[ʃ.ʃ.s] as deviant could not have been compared; as the violation point and phoneme was to be kept constant. With no significant phoneme effect, [s.s.s] as standard and *[s.s.ʃ] as deviant could be compared; a possible MMN to this comparison would show that the brain response stems from the harmonic vs. disharmonic distinction and not from the phoneme distinction.

The results in Figures 4.2 and 4.3 do not show a clear asymmetry between harmonic and disharmonic words; no MMN-like the response was observed. Even for the simple main effect (Figure 4.3), the topoplot shows no negative deflection during the latency range of MMN (100-300 ms) across the frontal-central region. For both the main effect comparison and the other single factor comparisons, no statistically significant effect was observed in the latency range of MMN component. The only statistically significant result was that the violation point affected the amplitudes. However, this effect is not interpretable by itself without a significant condition by violation point interaction. It only means that the amplitude was overall smaller when the violation point was in the middle.

These results were unexpected because behavioral (sensitivity index rate) results showed that participants were able to make a distinction between standards and deviants. The fact that no correlation was found between the behavioral and brain response implies that we were not able to measure the brain response indexing pattern violation. The question that arises here was why behavioral discrimination could not be observed in the perceptual system. One possible explanation for the lack of MMN effect relates to the stimuli that were used in the experiment. In particular, it is possible that the stimuli were heterogeneous enough that the auditory system could not form a memory trace.

The stimuli used in the study were heterogeneous in the sense that vowels were all different and the place of the /k/ phoneme differed from word to word. For example, out of 24 deviant words for *SH2-s, 14 words were in the form of [ʃ.s.k.ʃ] and 10 words were in the form [ʃ.k.s.ʃ]. Since we are dealing with very short time frames in milliseconds, the position of the /k/ is very important with respect to the MMN. Considering these two types of words as the same examples of disharmonic stimuli (both *SH2-s) may have caused a problem in the MMN analysis. In aggregate, to compare the standard with the deviant, the stimuli must be homogeneous in the sense that consonant /k/ must be the same place across standards and deviants. Moreover, when the responses to the onset of the word were analyzed, it was observed that some subjects didn't have clear AEPs.

In conclusion, even though the first analysis could not show that the auditory system was able to extract a repetition regularity out of complex linguistic stimuli, the problem discussed above cannot be ruled out. Therefore, Analysis 2 was conducted to clarify the issue by segmenting the stimuli in such a way that all the comparisons included homogeneous stimuli. To this aim, I excluded the disharmonic stimuli where /k/ was the third consonant – only the words in which /k/ was the second consonant were included. Also, only the participants who had robust AEPs to the onset of the word were included in the second analysis, which left 10 subjects.

4.3.2.3 Analysis II

EEG comparisons were made over combined cells, as discussed above, by averaging only the ten subjects. Figure 4.5 below displays the ERPs obtained for standard and deviant words as well as the difference waveforms for the main effect of condition (SH vs. *SH) irrespective of the violation point; and the topoplot which shows the mean difference wave at the horizontal line in the waveform plot at 200 ms. Both the waveform and the topoplot shows that there is positivity in the latency range of a MMN component, contrary to the expected negative deflection.

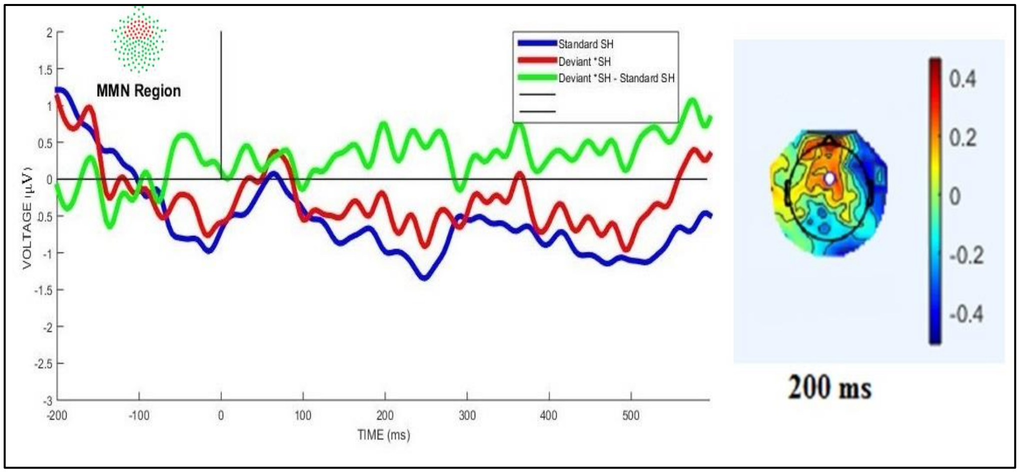


Figure 4.5: Time course of corresponding absolute voltage and difference waves in the grand average voltage for SH vs. *SH (left panel); and spatial distribution of difference wave at 200 ms (right panel).

This time ANOVAs were computed with four within-subject factors, as time was added as a within-subject factor with five levels (100-200 ms, 200-300 ms, 300-400 ms, 400-500 ms, 500-600 ms). The reason for this is to see whether there is a significant positive or negative deflection in a specific time interval during the last 500 ms. ANOVA results showed that the amplitude (μV) was not significantly affected by any of the factors or interactions, all p values >.05; though violation point again significantly affected the mean voltages, F(1,9)=25.541, p=.001. Also, there was a marginal main effect of the condition by violation point interaction, F(1,9)=4.341, p=.067.

4.3.2.4 Discussion of Analysis II

Although the data included only a subset of homogeneous stimuli and a subset of participants who showed robust AEP, Analysis II still didn’t show a statistically significant difference for the two different violation points by grammaticality and the main effect of grammaticality. When time as a five-level within-subject factor was added, a marginal main effect of the grammaticality by violation point interaction was observed. However, without a significant main effect of grammaticality, any significant interaction would be uninterpretable. All in all, no statistically significant result apart from the marginal significant effect of the grammaticality by violation point interaction effect was found.

4.4 General Discussion

By recording ERPs while manipulating and violating the participants’ auditory expectations at two distinct positions with two distinct phonemes, I aimed to obtain direct perceptual evidence which would show that an active and predictive system is being utilized to make predictions about lab-learned phonotactic rules.

We replicated the behavioral evidence for learning an attested rule with different test designs and showed that participants were able to make a behavioral distinction between standards and deviants (Experiment 1). If we assume that the rule was successfully acquired, there is a discrepancy between the learner’s behavior and their early auditory neural response. This suggests that while the pattern is learnable, and while it must exist in the learner’s grammar, we were not able to measure whether it is accessible to the auditory processor or not. On the contrary, several studies indicated that the MMN parameters correlate closely with the subject’s behaviorally determined perceptual accuracy (Kraus et al., 1996; Menning, Roberts, & Pantev, 2000). And the correlation between the MMN parameters and behavioral responses implies that pre-attentive processes determine the accuracy of the subsequent conscious processes (Tervaniemi et al., 2001). However, these results imply that the mechanism behind the perceptual sensitivity works differently than behavioral sensitivity. Nevertheless, it is hard to make an inference from null results; maybe, it is simply that we are unable to measure the neural activity associated with this type of deviance detection/prediction error using MMN.

My anticipated model for the auditory system is the following. Firstly, during the training phase, phonological pattern detectors would extract and acquire the abstract rule, and the rule would become available to the auditory processor. Then, depending on the consecutive standard stimuli, the auditory system generates a memory trace. The memory trace then enables the auditory processor to make predictions about incoming stimuli on the basis of whether they are abiding or violating the rule. In the final phase, there should be an observable neural reaction of surprise when the deviant differed from the prediction generated by the memory trace. However, the fact that no perceptual mismatch response was observed suggests that the auditory processor did not, in fact, generate such an abstract representation. Of course, there is another possibility that we simply could not measure it. Whether the auditory cortex had access to this rule knowledge is unclear, but it might not have used such knowledge to inform the creation of the memory trace (and did not use rule knowledge to make predictions about incoming sounds).

Since the rule is active at the behavioral level and participants used that rule to judge the grammaticality of the incoming stimuli, we can assume that the rule had been extracted. Consequently, the absence of a neural response can be explained in two ways: the rule in the grammatical system is inaccessible to the auditory processor, or the rule is accessible, but the auditory processor isn’t able to construct a representation based on the violation of the rule. In the former case, the rule is acquired and encapsulated in such a way that the auditory processor is not authorized to see the rule. In the latter case, the auditory processor is authorized to see the rule, but unable to use that information when creating a memory trace and generating predictions about incoming sounds. In either case, it would lead to the creation of a memory trace which does not contain any information about rule violation. Without this information indexed in the memory trace, it is impossible for the auditory processor to generate predictions about whether incoming sounds will follow the rule. To sum up, internal grammar features which govern language perception cannot be represented at the perceptual level.

Another explanation of the lack of neural correlate could be that the MMN is using a very limited time window (Wacongne et al., 2011). Wacongne et al. (2011) found a double dissociation between the early MMN and P3b which is a later temporally extended and distributed response. The study discusses that local violations of transition probabilities elicit MMN, but higher-order regularities cannot be captured by this reduced level response. For example, in a stimulus like [xxxxY], the last sound Y, which was a deviant from the standard [xxxxx], elicited a MMN due to generating a local deviancy. The sequence [xxxxY] systematically elicited a MMN even when the sequence itself is frequent and predictable On the contrary, when the sequence [xxxxY] was used as standard and violated on a rare subset of trials like [xxxxx], P3 was elicited due to the global deviancy even when the rare sequence was just a repeating of monotonous sounds (xxxxx). (Basirat, Dehaene, & Dehaene-Lambertz, 2014).

The kind of local deviancy is like an SL pattern (e.g., a spreading rule in phonology) in the subregular hierarchy. Contrariwise, the SH rule, which is an SP pattern, is a non-adjacent global rule and comprises a higher-order regularity. Wacongne et al. (2011) also found a late (∼300 ms) divergence (P3b) in response to this kind of global rule which reflected solely the deviance of the overall sequence rather than of its individual component tones. In this respect, Wacongne et al. (2011) view the MMN as a limited perceptual response which uses the recent past to predict the present and is blind to higher-order global rule violations. In other words, since the MMN is a limited perceptual response and does not have access to higher-order regularities, the mismatch response is not able to reflect internal grammar features. This implies that the early auditory processing stage is not sensitive to non-adjacent patterns and other ERP indexes such as P3b could be elicited. Nevertheless, I didn’t find a statistically significant positive peak of ∼300 ms.

I conclude that the detection of novelty in auditory language stimuli elicits a behavioral response. However, we were not able to measure the perceptual response (MMN) to novel stimuli after a training session, nor did we find a statistically significant response to a rule violation (like P3b). Two possibilities were discussed: either the MMN is blind to non-adjacent higher-order phonotactic rule violations, or I just failed to measure the appropriate neural response. The local-global paradigm literature (Basirat, Dehaene, & Dehaene-Lambertz, 2014; Bekinschtein et al., 2009; Wacongne et al., 2011) implies that the non-adjacent sibilant harmony pattern, a type of non-local fhigher-order rule, cannot be captured via MMN, which is blind to global rule violations. This indicates that the early auditory processing stage is not sensitive to violations of non-adjacent patterns.

4.5 Limitations and Directions for Future Research

The stimuli used in this study were taken from Lai (2012, 2015), and the assumptions and experimental design of Lai (2015) were different from the current EEG study. Although I tried to overcome these limitations by using homogeneously segmented data, I still did not observe a significant neural response to rule violations. Therefore, for future studies, I will try to control every possible artifact. Truckenbrodt, Steinberg, Jacobsen, & Jacobsen (2014) address some aspects of preparing stimuli for an ERP-experiment on speech processing and phonological learning. It has been suggested that lower-level acoustic stimulus characteristics should be controlled in order to avoid confounds with higher-level linguistic factors. Moreover, the stimuli should be kept as natural as possible and the artifacts caused by manipulation should be normalized. For the next experiment, therefore, the segmental durations of the stimuli across contrasts will be normalized and the sibilant distinction will be based only on spectral phonetic parameters.

Chapter 5

EXPERIMENT 3: CAN LPC INDEX PHONOTACTIC RULE VIOLATIONS?

After Experiment 2 failed to find the MMN to violations of non-adjacent phonotactic rule violations, I decided to instead run an experiment designed to elicit a Late Positive Component (LPC). Domahs et al. (2009) and Moore-Cantwell et al. (submitted) reported a higher amplitude Late Positive Component (LPC) to novel words that violated a learned phonotactic constraint than novel words that satisfied it. The LPC has been reported in response to violations of many kinds of long-distance sequencing rules, both syntactic and phonotactic. Thus, the aim of the current experiment was to observe an LPC to a phonotactic rule violation in the absence of semantic information, using a sibilant harmony rule which is an attested long-distance harmonic pattern.

After the first unsuccessful attempt in Experiment 2, the natural stimuli was transformed by controlling the duration of each phoneme. Since the aim is to partially replicate Moore-Cantwell et al. (submitted), the design of the test session was changed to match that of Moore-Cantwell et al. (submitted)'s test phase. It was a partial replication because a different pattern was used, namely the Sibilant Harmony. Also, the training phase was different from Moore-Cantwell et al. (submitted)'s training in which picture matching was used to teach new words. The training phase of this new experiment was the same as Experiment 1 (listen and repeat). In the test phase, which is not an auditory detection task anymore, participants rated the likelihood that each word followed the rule they had learned. Thus, it was a grammaticality judgment task. One-third of the test words were present during training, one-third of the words were novel and sibilant-harmonic, and the other one third were novel and disharmonic. The details of the methodology are outlined in section 5.1 below.

Behavioral results showed that there was a significant mean rating difference between novel-harmonic and disharmonic words. The stimuli elicited a robust auditory evoked potential (AEP) to the word onsets. However, no LPC was observed following the stimulus or response, and there was no significant difference in ERP response to harmonic and disharmonic words. Both behavioral and EEG results are presented in section 5.2. Possible reasons behind these results are discussed in section 5.3, and this chapter ends with a brief conclusion.

5.1 Method

5.1.1 Participants

A total of 18 University of Delaware students were recruited as participants and provided written consent in the experiment. 3 participants were eliminated because of excessive artifacts (with less than 20% good trials), yielding a final sample of 15 subjects. Each participant received course credit for participation. 14 of the 15 participants were females and 1 was male (this imbalance arises from the fact that the population we convenience-sampled from, undergraduates in University of Delaware Linguistics Department, was overrepresented with women). All participants provided informed consent to participate. Two participants were left-handed, but I did not exclude left-handers, as most left-handed people have left-lateralized language function

(Somers et al., 2015). The mean age was 19.4 (*SD = 1.05, range = 18 to 21*). None of the participants reported a history of hearing loss or speech/language impairments, and all reported having English as their first and only language.

5.1.2 Stimuli

The sibilant harmony rule (sibilants in a well-formed word have to agree in anteriority) as a phonotactic pattern was chosen to be learned by the subjects. The stimuli used in this study was adapted from Lai (2015), but I used only a subset of stimuli that were homogeneous in the sense that the place of consonants in the word was fixed across all the stimuli. All training and test stimuli consisted of three syllables of the form of CV.kV.CVC. All the stimuli had a stop [k] as the second consonant with sibilants ([s, ʃ]) as the first, third, and last consonants. The vowels in the alphabet of the language were [a, ɛ, ɔ, i, u]. 80 words were constructed with the sibilants in 60 them agreeing in anteriority (e.g. [sakesos], [ʃekoʃiʃ]), and in 20 disagreeing (e.g. [sekaʃos], [ʃukoseʃ]). Of the 60 words that had an agreement, 40 appeared only in the training session. The other 20 were used in the test phase as novel-fit items (together with 20 familiar words and 20 novel-violate words). Table 5.1 below summarizes the types of training and test stimuli used.

Table 5.1: Harmonic and disharmonic stimuli used in both training and test phase

Sibilant Tier	Harmonic	Disharmonic
[s.s.s]	[sakesos]	None
[ʃ.ʃ.ʃ]	[ʃekoʃiʃ]	None
[s.ʃ.s]	None	[sekaʃos]
[ʃ.s.ʃ]	None	[ʃukoseʃ]

The words were pronounced in CV.CV.CVC format (Lai 2012, 2015). In order to control the duration of the words, the first, second and third syllables were segmented from the sample words using the offset/onset of noise for the surrounding vowels as a criterion. I then applied a 10 ms sinusoidal fade in and fade out to the beginning and end of each recording to eliminate the effects of trimmed formants. Next, I merged three syllables and constructed the transformed CV.kV.CVC words. As a final step, the peak amplitude of all items was normalized to their mean. See Figure 5.1 for a spectrogram view from Praat software (Boersma & Weenink, 2017).

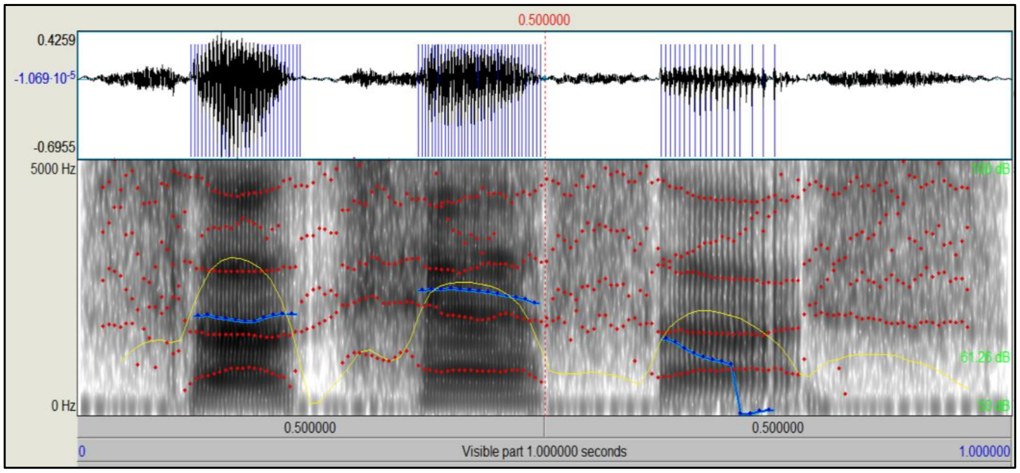


Figure 5.1: Praat spectrogram for an example word, [sakesas].

Segmental durations were equated by setting the initial syllable (e.g., [sa] or [ʃa]) to 250 ms, the second syllable (e.g., [ka]) to 250 ms and the third syllable (e.g., [sas] or [ʃaʃ]) to 500 ms. This makes the duration of all the words 1000 ms (whether it is following the rule or violating the rule) in which the first sibilant starts at 0 ms, second sibilant at 500 ms (making the violation point (in novel-violate words) beginning at 500 ms), and the third sibilant at 775 ms. All of these operations were done by using Praat software (Boersma & Weenink, 2017).

5.1.3 Apparatus and procedure

The apparatus was the same as the previous experiments. The procedure for the experiment consisted of two phases: a training phase and a testing phase. Participants, after reading the instructions, started a training block by pressing any button on a response pad. The training trial began with the appearance of a fixation cross on the computer monitor. One of the 40 training words conforming to the harmony pattern was presented from the loudspeakers 750, 1000, or 1250 ms later (randomly selected, jittered duration of fixation). The fixation cross remained on the screen for 500 ms after the offset of the word and was then replaced by a <repeat the word> prompt. Once participants had repeated a word, they pressed any button to proceed to the next word. EEG data collected while the participants were repeating the words was NOT analyzed, and participants were free to move their eyes and blink during these presentations. In each training block, all 40 words were presented 5 times each in random order (200 trials per block). Duration for one training block was approximately 15 minutes. The training phase was an exact replication of Lai (2015) with no meaning associated with the words.

During testing, following Moore-Cantwell et al. (submitted), participants were informed that some of the test words had been heard during training and were clearly part of the language. They were asked to rate how likely it was that each word was part of the language they were learning. They were encouraged to use all four response buttons rather than only the 1 labeled “unlikely a word” and the 4 labeled “very likely a word.” Participants began each test block by pressing any button. At the beginning of each trial, the fixation cross was shown on the computer monitor. One of the test words was presented over the loudspeakers 750, 1000, or 1250 ms (randomly selected, jittered duration of fixation) later. The fixation cross remained on the screen for 1000 ms after the word offset and was then replaced by the response prompt <Likely a word?> with the labeled scale. A test trial ended after a response was given. In each test block, all 60 words (twenty each of Familiar, Novel-Fit, and Novel-Violate) were presented once in random order. Duration for one test block was approximately 5 minutes.

The training block – test block sequence was repeated 3 times for a total of 600 training trials and 180 test trials (60 of each type). For all trials, when the fixation cross was shown on the screen (when participants are listening to the words) participants were asked to refrain from blinking, moving their eyes, or moving any other part of their body, including moving a finger to press a button. They were encouraged to make these and any other movements while repeating the words (during training) or looking to the response prompt (during testing). Participants were not allowed to take a break between a training and test block; they were encouraged to take breaks after each test block. At the end of the experiment, participants were asked if they had noticed anything about the language they had learned and if they had developed any strategy to distinguish between words that were and were not in the language. The total duration for both training and testing was about 60 minutes.

5.1.4 Data recording and analysis

The behavioral measure in this experiment was the mean ratings; during testing, participants rated the words from 1 to 4 indicating how likely they were to be in the language that they were learning. Mean values were compared in a repeated-measures analysis of variance (ANOVA) with the two factors: Block (three levels: 1,2,3) and Word Type (three levels: Familiar, Novel-Fit, and Novel-Violate).

Electrophysiological recordings were done in the same way as in Experiment 2A. After recording, the continuous EEG was segmented into epochs of 1200 ms for stimulus-locked averages and 1000ms for violation-locked averages. The baseline period was -200 to 0 ms for stimulus-locked and -400 to 0 ms for response-locked averages. A 0.1 – 40 Hz bandpass filter was applied prior to latency measures. Artifact detection and bad channel replacement steps were also the same as the Experiment 2A. The signals were re-referenced offline to the average of the left and right mastoids. EEG from test trials was averaged by condition (Familiar, Novel-Fit, Novel-Violate) across all blocks.

The main ERP measurements were taken from absolute waves. It was measured at frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) electrode sites. For statistical analysis, amplitudes were measured as the mean voltage in a given measurement window. The time window for the LPC was 600 to 1000 ms after word onset at parietal electrode sites. An analysis of variance (ANOVA) was used for the LPC analyses that included word type (Familiar, Novel-fit, Novel-violate) in a one-way repeated measures design. And for the violation-locked analysis, time was added as a within-subject factor with six levels (0-100 ms, 100-200 ms, 200-300 ms, 300-400 ms, 400-500 ms, 500-600 ms) to check for an effect interacting with time. The reason for this is to see whether there is a significant positive or negative deflection in a specific time interval after the violation point. All significant ANOVA effects are reported with the partial η^2 effect size measure and the t-tests with Cohen's *d*.

5.2 Results

5.2.1 Behavioral Results

Figure 5.2 below shows the comparison of mean ratings of different words across blocks.

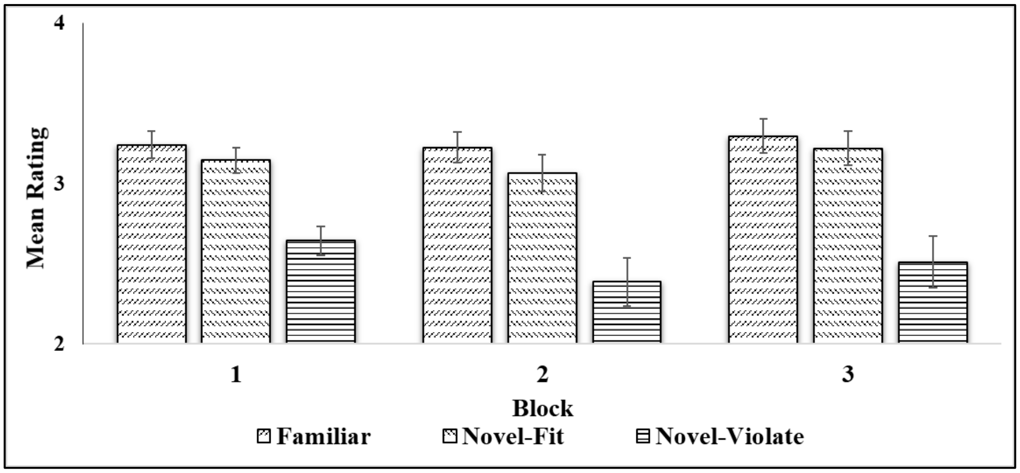


Figure 5.2: Mean ratings of Familiar, Novel-Fit and Novel-Violate word types across blocks.

ANOVA results showed that mean ratings were significantly affected by the Block, F(2,28)= 4.380, p=.022, η2=.238, 1-β=0.709; and significantly affected by the Word type, F(2,28)= 15.697, p<.001, η2=.529, 1-β=0.999.

However, there was no interaction between the two (p>.05). Pair-wise comparisons revealed significant differences in ratings of novel-fit words vs. novel-violate words (*MD=.631, SE=.171*), t(14)=3.68, p=.007; and familiar words vs. novel-violate words (*MD=.742, SE=.174*), t(14)=4.265, p=.002. These results clearly show that participants learned the pattern and rated Novel-violate words as outside of the language in all three blocks.

5.2.2 Electrophysiological Results

5.2.2.1 Stimulus-locked

Figure 5.3 below shows the time course of corresponding absolute voltage for the Novel-fit vs Novel Violate word types across the midline electrode regions. Although there was a very clear AEP to the onset of the word, the comparison of ERPs elicited by Novel-fit and Novel-violate in the late time window of 600 ms to 1000 ms after the word onset revealed no significant difference; (F(2,28)= .759, p>.05).

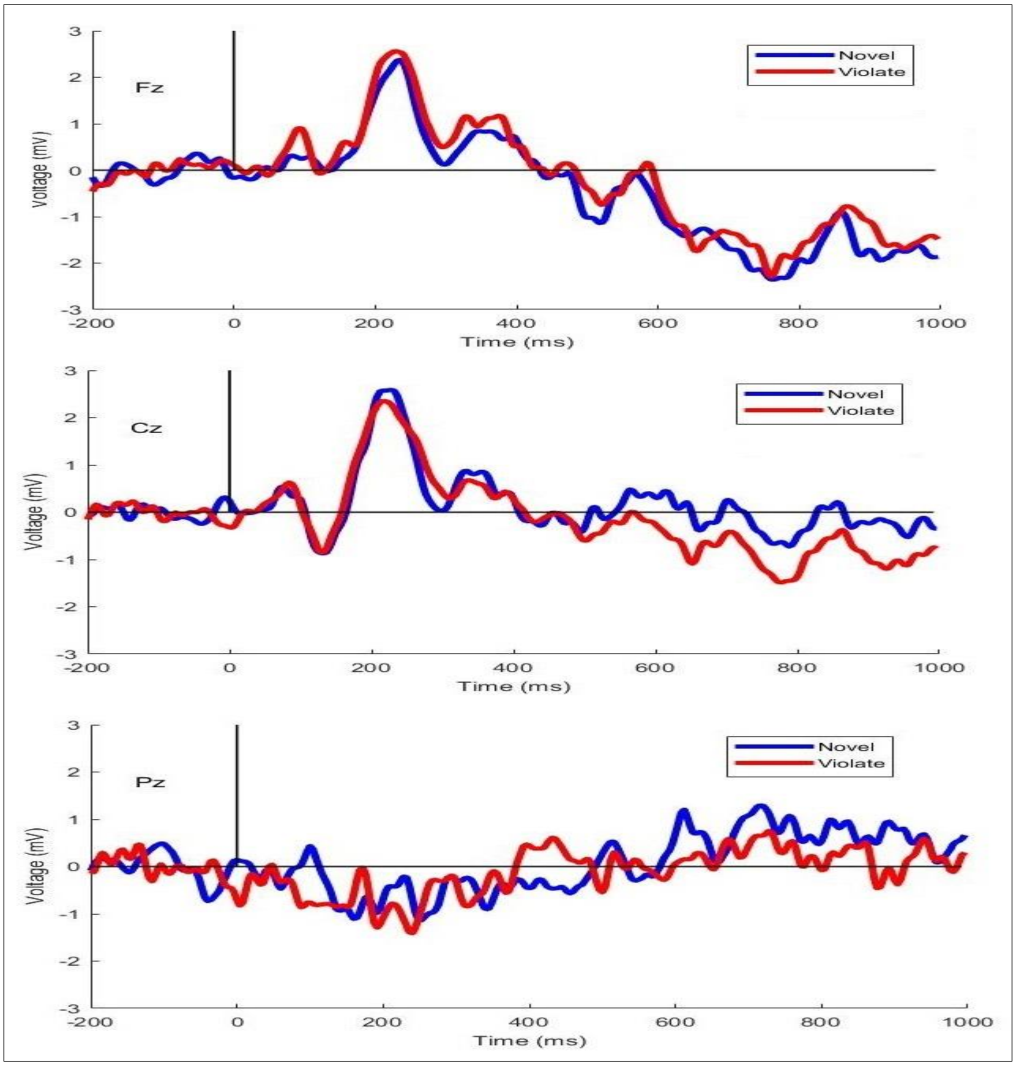


Figure 5.3: Time course of the corresponding absolute voltage for the Novel-fit vs Novel Violate word types across midline electrode regions.

5.2.2.2 Violation-locked

Figure 5.4 below shows the time course of violation-locked absolute voltage for the Novel-fit vs Novel Violate word types across the midline electrode regions. The aim was to see if there was a significant difference between the two waveforms after the violation point by looking at every 100 ms interval.

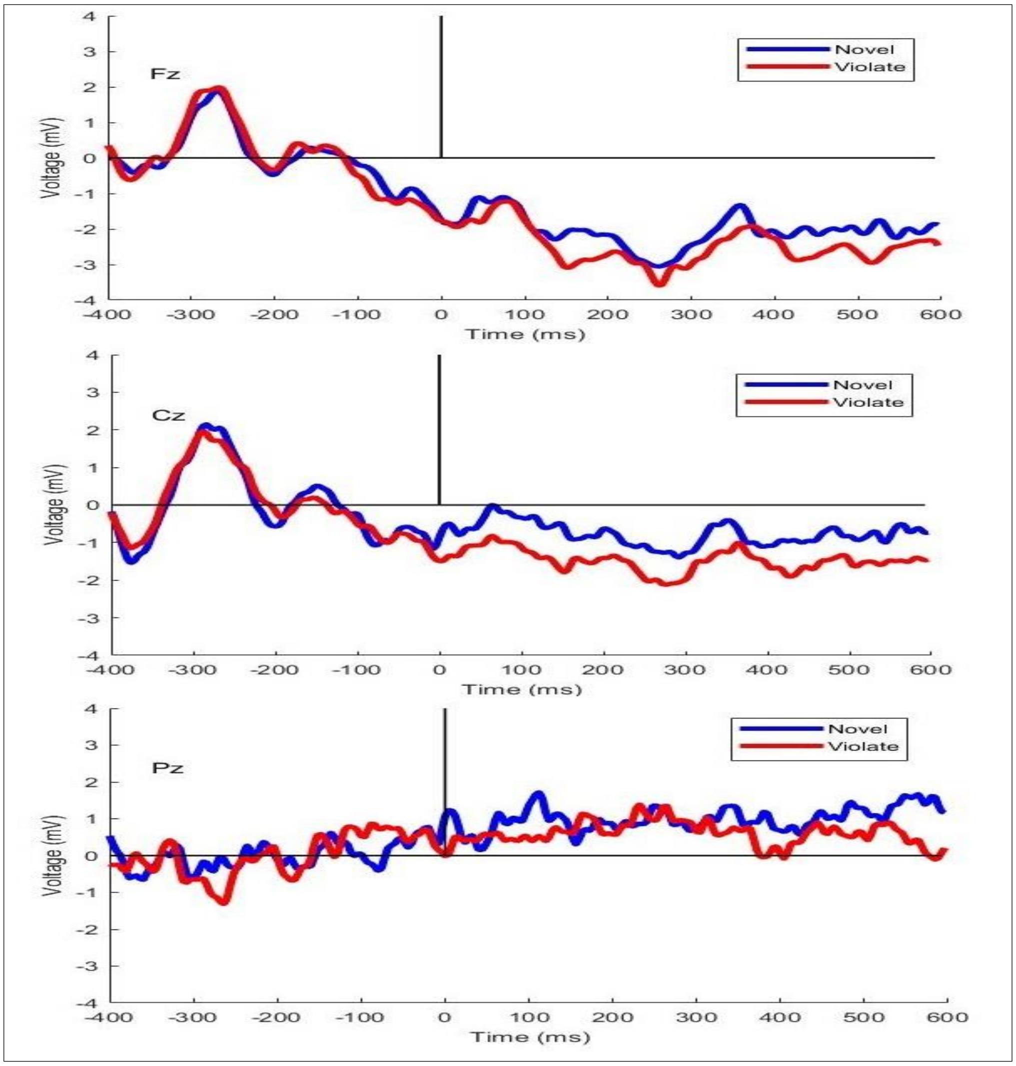


Figure 5.4: Time course of corresponding absolute voltage for the Novel-fit vs Novel violate word types across midline electrode regions.

However, the comparison of ERPs elicited by Novel-fit and Novel-violate after the violation point revealed no significant difference; there was no word type effect (F(2,28)= 1.097, p>.05) and no time interval effect (F(5,70)= 1.659, p>.05).

5.3 Discussion

In this experiment, the aim was to observe an LPC in response to a phonotactic rule violation. The stimuli used in this experiment were in the *C.V.C.V.C.V.C* format and strictly controlled in order to control for phonetic differences and duration. The behavioral results show that participants were able to make a distinction between novel-fit and novel-violate words after the training session. However, EEG results showed that the LPC was not elicited in response to novel words violating the phonotactic pattern. In fact, a late positive component was entirely absent from the data, even though, a robust AEP with P1, N1, and P2 peaks was observed.

This experiment partially replicated the Moore-Cantwell et al. (submitted) study (henceforth MC study) where the LPC was shown to index phonotactic rule violations. Our experiment differed from the MC study in two respects: (i) training session, (ii) length of the stimuli. In the training session of the MC study, participants learned words of the made-up language by matching the spoken words to pictures of the objects the words refer to and were given feedback. Compared to our training session, where subjects only listen and repeat the words, this study incorporated semantic information in the learning. By doing this, the training phase was supposed to mimic a natural word learning scenario where infants use the word to world mapping (Booth & Waxman, 2003). One of the reasons the LPC was observed in the MC study but not in our study could be due to the differences between the training sessions.

Second, in the MC study, the words were in the CV.CV format with a duration of averagely 367 ms where the 2nd syllable began 136-245 ms after the first. In our partial replication, the duration of words was fixed to 1000 ms where the violation was at 500 ms. This shows that the length of the stimuli was very different. Although both patterns (Sibilant Harmony rule and Devoicing rule) are non-local at the segment level, the pattern that I tested in this experiment was much more complex in terms of the number of sibilants and the long-distance dependency. Therefore, the computational complexity of the stimuli in the current experiment could be the other reason for not observing LPC.

5.4 Conclusion

In this chapter, an EEG experiment that was designed to correlate violations of phonotactic rule learning and the neural response was reported. We tried to find an LPC in response to anomaly detection in a rule-governed phonological sequence. However, no LPC was observed. We speculate that the reason for this LPC absence was related to the computational complexity of the stimuli used. Therefore, for the next EEG experiment, I simplified the stimuli and used words in the CV.CV format which would still be a non-adjacent dependency at the segment level but adjacent at the syllable level.

Chapter 6

EXPERIMENT 4: NEURAL MARKERS OF COMPUTATIONALLY SIMPLER PATTERNS

In this chapter, the results of the fourth experiment that was conducted to observe neurophysiological markers of phonological rule learning with simpler stimuli will be presented. After the two null results EEG experiments that were presented in Chapters 4 and 5, we decided to simplify the rule by using computationally simpler stimuli and incorporate the rule knowledge in a categorization task and look for a P300 as an index of categorization between harmonic and disharmonic words. In addition, we looked for an LPC as an index of anomaly detection in a rule-governed phonological sequence. After the simplification of the pattern, the aim of this experiment was more focused on finding the neurophysiological markers of the implicit phonological rule learning rather than comparing the neural markers of attested vs unattested patterns. In terms of neurophysiology, abstract rule learning should translate into neurally observable prediction mechanisms, as in Kuhl’s Native Language Neural Commitment (NLNC) Hypothesis (Kuhl et al., 2008). This means that if a rule is learned after exposure to repeated stimuli, this learning should lead to an instant, measurable "tuning" of the perceptual system. The aim was to test this perceptual tuning in this chapter.

A controlled set of stimuli that followed a very simple grammar consisting of a sibilant harmony rule was constructed. After a brief training session, participants were asked to categorize each stimulus according to the grammar they learned while their brain responses were measured with EEG. In the testing phase, the task for the participant was to press a button in response to each stimulus to categorize the stimulus as grammatical or ungrammatical, according to the rule they had learned in the training phase. Participants responded with the left hand for grammatical words and with the right hand for ungrammatical words. Participants were tested in an auditory oddball paradigm and received one block in which a deviant (ungrammatical) stimulus appeared infrequently among repeated occurrences of a standard (grammatical) stimulus, and one block in which the frequency was reversed. By changing the frequency across the blocks, the goal was to see whether the rule knowledge persists even when participants are presented with more ungrammatical stimuli. The order of these two blocks was not randomized as receiving frequent ungrammatical stimuli right after the training session would not make sense. Specifically, exposure to frequent ungrammatical words could erase participants rule knowledge right after the training session. The experimental details are explained in section 6.2.

The results, presented in section 6.3, showed that participants learned the simple rule at the behavioral level (as measured with d-prime, a sensitivity measure to rule violations). In both blocks, participants showed behavioral sensitivity independent of the frequency of stimulus presentation. However, this rule learning is reflected in the brain response to violations of the rule only in Block I. While grammatical words were frequent (Block I), rule learning was reflected in both a P3 difference waveform and an LPC, whereas when ungrammatical words were presented frequently, the difference at the neural level was lost. I discuss the implications of these results in section 6.4, and section 6.5 concludes the chapter by reporting that brief sessions of laboratory learning of phonotactic rules result in a neural commitment which is reflected in the brain response.

6.1 Recap

The particular brain response elicited in an ERP study depends on several factors, such as the design features of the experiment (attentive or inattentive), the nature of the stimuli (linguistic or non-linguistic), and even the complexity of the pattern or rules (local or higher-order). Processing of phonological violations has been shown to elicit different kinds of electrophysiological responses including MMN, P3, and LPC. As pointed out previously, Aaltonen et al. (2008) reported a mismatch negativity (MMN) in response to native language violations of non-adjacent phonological constraints. However, there is a rich literature demonstrating that the MMN is a limited perceptual response which uses the recent past to predict the present and is blind to higher-order global rule violations (Basirat, Dehaene, & Dehaene-Lambertz, 2014b; Bekinschtein et al., 2009; Chennu et al., 2013; Tse, Low, Fabiani, & Gratton, 2012; Wacongne et al., 2011). Using a hierarchical “local-global” paradigm, it has been shown that novelty detection in auditory stimulation includes two different neural components: (i) the MMN reflecting an unconscious pre-attentive response, and (ii) the P3 indexing conscious access to working memory updates. Local violations of transition probabilities elicit a MMN, while higher-order regularities (global rules or non-local patterns) which solely reflect the deviance of the overall sequence rather than of its individual components are captured by the P3[6], a positive-going wave peaking at around 300 ms. The P3 latency reflects stimulus evaluation and encompasses the perception and categorization of the stimulus for the purpose of making a behavioral response (for more information about P3-like responses, see Polich, 2007). Moreover, McLaughlin et al. (2010), Domahs et al. (2009) and Moore-Cantwell et al. (submitted) reported an LPC to violations of phonotactic rules.

The aim of this experiment was to look for P300 as an index of categorization between harmonic and disharmonic stimuli in addition to the LPC as an index of anomaly detection in a rule-governed phonological sequence. The current experiment was designed to use the P3 wave to determine whether the categorization of grammatical vs ungrammatical words (according to a lab-learned artificial grammar) is indexed by this ERP component within a specified time window. The difference wave was used to isolate the cognitive processes related to deviance detection. If the amplitude of the rare-minus-frequent difference wave is different from zero, it can be concluded that the process before and during the categorization is different for grammatical words compared to ungrammatical words (see Luck et al., 2009, for the details of using P3 difference waveform to isolate related neural effect).

---

[6] Specifically, it is P3b, which is a task-relevant potential elicited during target stimulus processing (Polich, 2007),

6.2 Method

6.2.1 Participants

A total of 28 University of Delaware students were recruited as participants and provided written consent for participation in the experiment. 4 participants were eliminated due to exceptionally noisy EEG data, yielding a final sample of 24 participants. Each participant received course credit for participation. 22 of the 24 participants were female (this imbalance arises from the fact that the population I sampled from was overrepresented with women). One participant was left-handed, but I did not exclude left-handers, as most left-handed people have left-lateralized language function (Szaflarski et al., 2011). The mean age was 19.87 (*SD=1.68*, *range=18 to 24*). None of the participants reported a history of hearing loss or speech/language impairments, and all reported speaking English as their first and only language.

6.2.2 Stimuli

A simplified version of the Sibilant Harmony rule was used as the phonotactic pattern to be learned by the participants. The stimuli used in this study again were adapted from Lai (2012, 2015). The stimuli were originally recorded from a native speaker of Mandarin Chinese with phonetic training, pronouncing the words in CV.CV.CVC format. The third syllable without coda (CV) was segmented from the sample words using the offset/onset of noise for the surrounding vowel and sibilant as a criterion. I then applied a 10 ms sinusoidal fade-in and fade-out to the beginning and end of each recording to eliminate the effects of trimmed formants. Next, I merged two syllables and constructed the CV.CV words. As a final step, the peak amplitude of all items was normalized to their mean. See Figure 6.1 for a spectrogram view from Praat software (Boersma & Weenink, 2017). The duration of each phoneme was strictly controlled at 100 ms, making each word 400 ms long, the second syllable beginning at 200 ms.

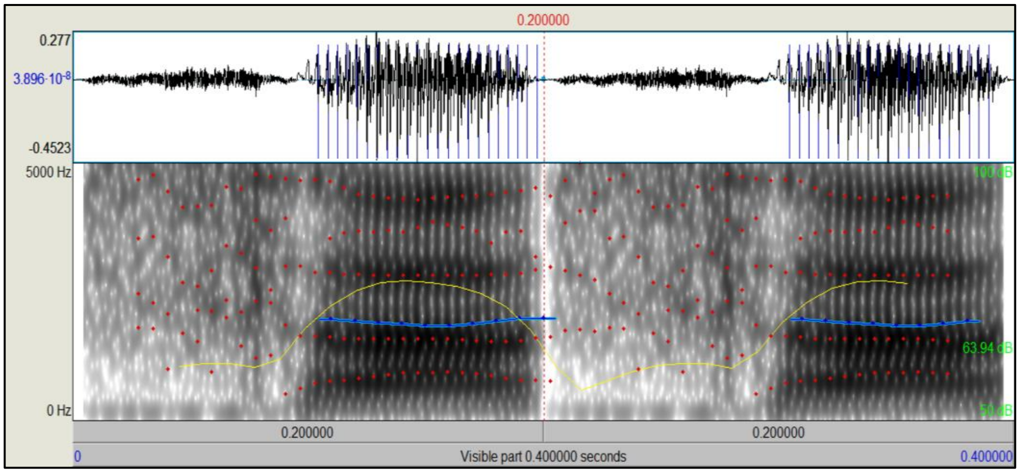


Figure 6.1: Praat spectrogram for an example word, [sasa].

All training and test stimuli consisted of two syllables of the form CV.CV, with sibilants ([s, ʃ]) as the first and second consonants. The vowels in the alphabet of the language were [a, ɛ, ɔ, i, u]. 100 words were constructed with half agreeing in sibilant anteriority (e.g. [saso], [ʃeʃi]), and the other half disagreeing (e.g. [saʃi], [ʃeso]). Of the 50 words that had an agreement, 20 appeared only in the training session. The other 30 were used in the test phase as novel-fit items (together with 20 trained words). All of the 50 words that had a disagreement were used in the test as novel-violate items.

In order to exclude potential artifacts due to phonetic differences, as exemplified in Domahs et al. (2009), I checked the mean fundamental frequency (Hz) and mean intensity (dB). See Table 6.1 below for the statistical comparisons. It can be seen that with respect to phonetic parameters, there is no difference between grammatical vs. ungrammatical words.

Table 6.1: Phonetic properties of grammatical vs ungrammatical words (mean fundamental frequency and mean intensity) and comparisons between types of words.

Word Type	Sibilant Tier	F_0 (mean)	Intensity (mean)
Grammatical	[s.s] or [ʃ.ʃ]	240.3 (*SD=5.3*) Hz	71.6 (*SD=0.95*) dB
Ungrammatical	[s.ʃ] or [ʃ.s]	240.3 (*SD=5.2*) Hz	71.6 (*SD=1.02*) dB
Grammatical vs Ungrammatical		t(49)=.010; p=.992	t(49)=.107; p=.916

6.2.3 Apparatus and procedure

The apparatus was the same as the previous experiments. The procedure for the experiment consisted of two phases: a training phase and a testing phase. During the training phase, participants listened to 200 tokens (20 grammatical words repeated ten times) and were instructed to repeat each word orally after they heard it. The training lasted approximately 15 minutes. In the testing phase, the task for the participant was to press a button in response to each stimulus to categorize the stimulus as grammatical or ungrammatical, according to the rule they had learned in the training phase. Participants responded with the left hand for grammatical words and with the right hand for ungrammatical words.

Participants were tested in an auditory oddball paradigm and received one block in which a deviant (ungrammatical) stimulus appeared infrequently among repeated occurrences of a standard (grammatical) stimulus, and one block in which the frequency was reversed. The order of these two blocks was not randomized as receiving frequent ungrammatical stimuli right after the training session would erase rule knowledge.

Participants discriminated between grammatical and ungrammatical words, with grammatical words appearing in 83% of trials (250 words) and ungrammatical words appearing in 17% of trials (50 words) in the first block, and with ungrammatical words appearing in 83% of trials and grammatical words appearing in 17% of trials in the second block, with 600 total trials. Each participant was presented with a total of 100 rare stimuli and 500 frequent stimuli. The stimuli were delivered continuously, with a random number (between 3 and 7) of standards between each deviant. Each word was presented for 400 ms followed by a blank inter-trial interval of 1100 to 1500 ms. The duration of the test session was approximately 20 minutes. No explicit feedback was given to the participants during the test phase. All participants easily understood the instructions.

6.2.4 Behavioral data recording and analysis

In the test phase, button presses made by participants in response to deviant stimuli (ungrammatical words) were recorded. Hits (when an ungrammatical word was present, and the participant detected it and reported hearing it) and false alarms (when a grammatical word was present, but the participant reported hearing an ungrammatical word) were counted. The sensitivity index *d'* was then derived from the hit and false alarm rates according to signal detection theory (Macmillan & Creelman, 2004). When participants cannot discriminate at all, hits would be equal to false alarms, which would give a $d' = 0$. Results higher than zero mean that participants detected the signal. In the context of our study, positive *d'* means the rule was learned (in the sense that there is sensitivity to ill-formed words). *d'* was compared to zero – which would indicate no learning – by using one-sample t-test. Mean accuracy (ACC) and reaction times (RT) were also analyzed using paired samples t-tests comparing grammatical vs ungrammatical words separately for each block.

6.2.5 Electrophysiological data recording and analysis

Electrophysiological data were recorded continuously with 24-bit digitization at 250 Hz from 128 carbon fiber core/silver-coated electrodes in an elastic electrode net (Geodesic Hydrocel 128). The acquisition and digitization of the data were performed with EGI Net Station software v.4.5. Electrode impedances were lowered to below 50 kΩ before data collection. Each electrode was referenced to Cz. After recording, the continuous EEG was passed through a 0.1-40 Hz bandpass filter. The continuous EEG data were then segmented into epochs of 1200 ms for stimulus-locked averages and 1500 ms for response-locked averages. The baseline period was -200 to 0 ms for stimulus-locked and -1000 to -800 ms for response-locked averages. The next step was artifact detection in which bad channels, eye blinks, and eye movements were identified: if the difference between the maximum and minimum voltage exceeded 200 μV in a moving average of 80 ms, the channel was marked bad. If a channel was bad in over 20% of trials, it is considered bad in all trials; and trials containing more than 10 bad channels were marked as bad (Dien, 2010). A spherical spline interpolation was used to replace bad channels. Each trial was then baseline corrected using the mean voltage of the 200 ms pre-stimulus period.

The signals were re-referenced offline to the average of the left and right mastoids. Trials with incorrect behavioral responses were excluded from the averages (Woodman and Luck, 2003; Keil et al., 2014). The main P3 measurements were taken from rare-minus-frequent difference waves. The LPC was measured from absolute waveforms obtained for grammatical vs ungrammatical words. The P3 was measured at frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) electrode sites, and the LPC at a posterior region that includes sixteen channels determined by visual inspection.

Amplitudes were measured as the mean voltage in a given measurement window: 400 to 700 ms for the stimulus-locked P3, and -200 to -100 ms for the response-locked P3, following Luck et al. (2009). As for the LPC, the measurement window was from 600 to 1000 ms. Analysis of variance (ANOVA) was used for the P3 analyses that included factors of the region (frontal, central, parietal), and grammaticality (grammatical, ungrammatical words) for each block separately. LPC amplitudes were analyzed using paired samples t-tests comparing grammatical vs ungrammatical words for each block. All significant ANOVA effects are reported with the partial η^2 effect size and power (1-β err prob) measure and the t-tests with Cohen’s *d*. Correlations between behavioral measures (d', ACC and RT) and ERP measures (P3 and LPC) were assessed using the Pearson r correlation coefficient.

6.3 Results

6.3.1 Behavioral Results

In Block I, ungrammatical words were detected with a mean sensitivity of 0.557 (*d'*) (*SD=.82*), a score significantly greater than zero, t(23)=3.35, p=0.003, d=0.684, 1-β=0.894. Mean bias rate was -0.291 (c) (*SD=.31*), a value negative and significantly lower than zero which is the artifact of the oddball paradigm. See Figure 6.2 below for a visual comparison. This result indicates that the deviants were detected with a mean sensitivity higher than zero. Sensitivity for ungrammatical words was higher than zero sensitivity, showing that participants learned the rule. Mean accuracy (percent correct) for grammatical words was .70 (*SD=.14*) and .48 (*SD=0.19*) for ungrammatical. This difference was significant t(23)=4.78, p<0.01, d=.976, 1-β=.999. The accuracy data suggest only about half of the time the ungrammatical items were identified as ungrammatical. The above chance d' is likely to be driven by the high rate of correct-rejection of the grammatical words (assuming ungrammatical is our target). Mean reaction time was 545 ms (*SD=106*) for grammatical words and 564 ms (*SD=119*) for ungrammatical. This difference was also significant, t(23)=2.349, p=0.028, d=.479, 1-β=.736).

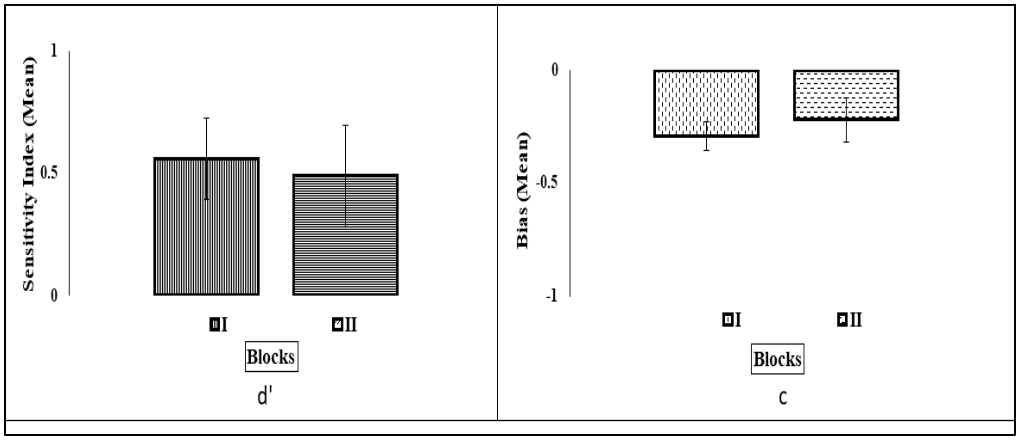


Figure 6.2: Block averages of sensitivity index rates (left panel) and bias rates (right panel). Error bars indicate standard error of the mean. No significant difference was observed between the blocks.

In Block II, ungrammatical words were detected with a mean sensitivity of 0.489 (*d'*) (*SD=1.026*), a score significantly different from zero, t(23)=2.33, p=0.029, d=0.477, 1-β=0.733. Mean bias rate was -0.220 (c) (*SD=.47*) and significantly different than zero (see Figure 6.2 above). These results indicate that sensitivity for ungrammatical words was higher than zero, showing that participants learned the rule. Mean accuracy (percent correct) for grammatical words was .67 (*SD=.16*) and .49 (*SD=0.25*) for ungrammatical. Again, the accuracy data suggest only about half of the time the ungrammatical items were identified as ungrammatical. The difference was significant t(23)=2.96, p<0.007, d=.604, 1-β=.889. Mean reaction time was 515 ms (*SD=147*) for grammatical words and 511 ms (*SD=146*) for ungrammatical. However, this difference was NOT significant, t(23)=0.261, p>0.05. Overall, behavioral results between the two blocks imply that the frequency of the stimuli does not have an effect on the behavior.

6.3.2 P3 Rare-minus-frequent difference waveforms

6.3.2.1 Stimulus-locked waveforms

Figure 6.3 shows stimulus-locked ERP absolute waveforms and the difference waves (rare-minus-frequent) across the blocks. It can be seen that both frequent and rare stimuli revealed a typical auditory evoked potential (AEP) with an N1–P2 wave-form complex at frontal-central electrodes in both blocks. It was important to get this AEP to the onset of the word whether the word was grammatical or ungrammatical. This indicates that the auditory system has responded to the onset of the word.

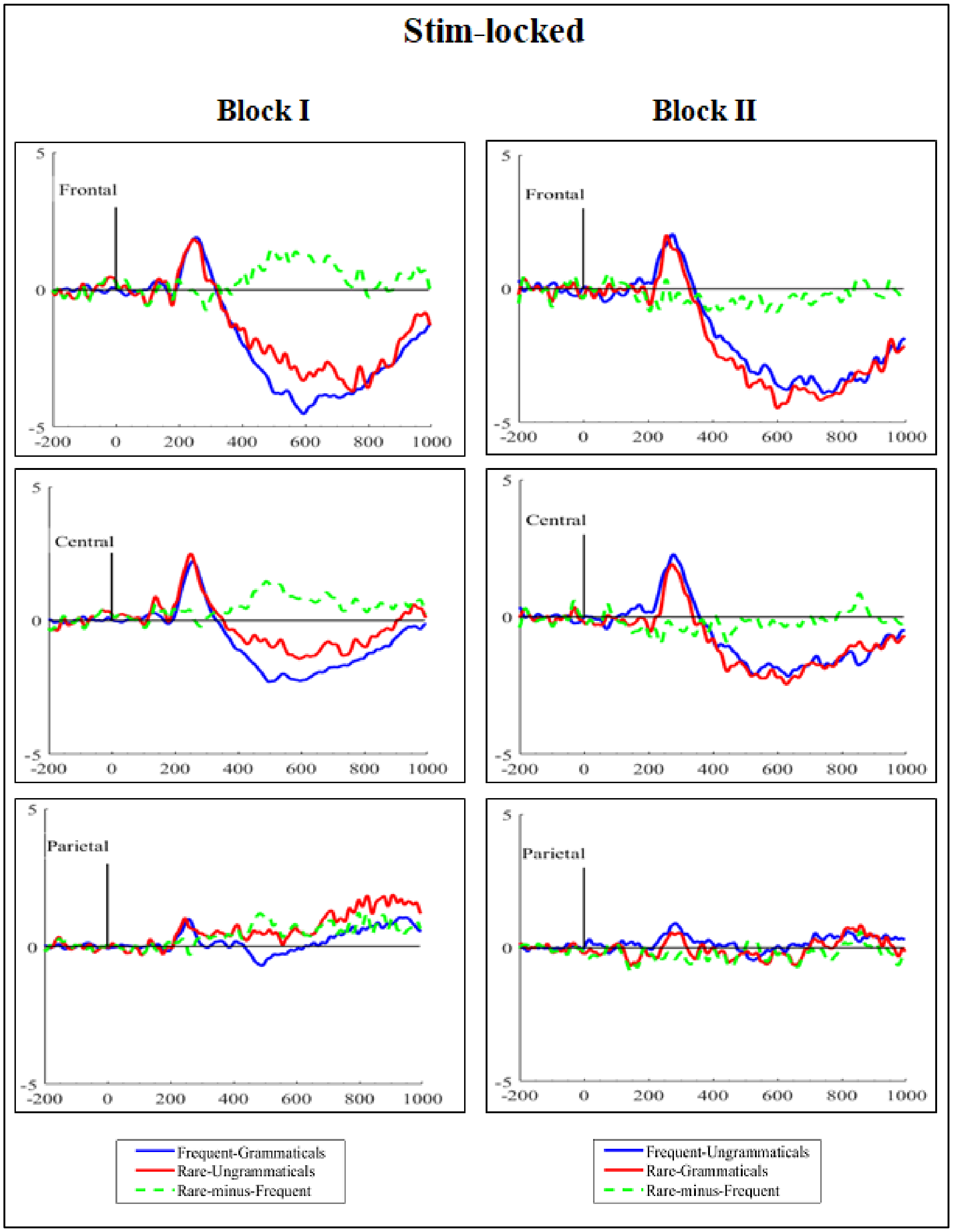


Figure 6.3: Stimulus-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites across blocks. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV).

In Block I, it can be seen that there is a noticeable difference between grammatical (blue) and ungrammatical (red) words beginning from the offset of the word (400ms) to 700 ms, specifically at the frontal and central regions. The amplitude of the difference wave was higher in the frontal region and lesser in the parietal region. The time course of the difference wave shows that the brain detected the violation at around 200 ms. This was marked by a P3 peak at 500 ms (~300 ms after the violation point). Whereas in Block II, both grammatical and ungrammatical words are processed the same; thus, the difference waveform was not significantly different from zero. A whole scalp snapshot of the effect at 500 ms can also be viewed across the blocks in Figure 6.4.

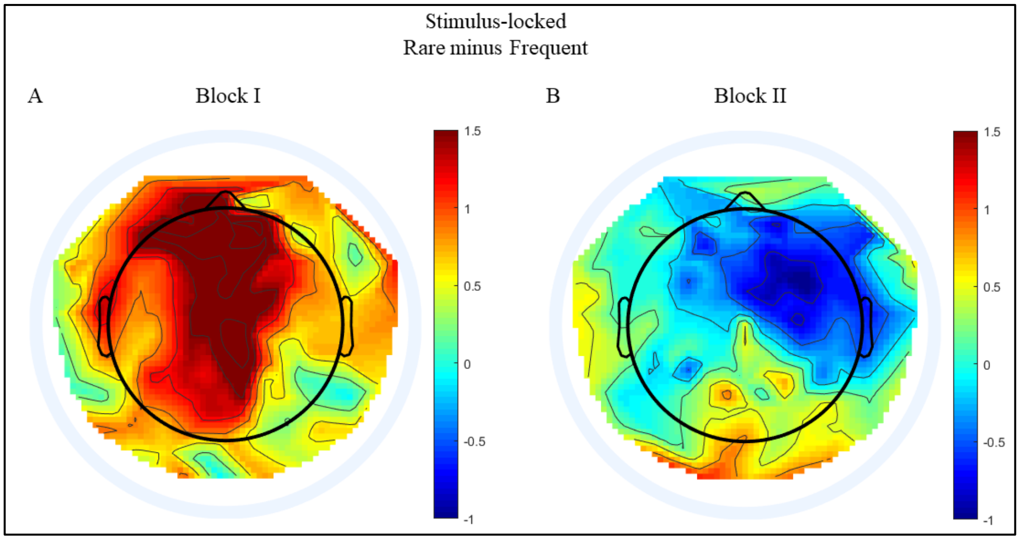


Figure 6.4: Spatial distribution of the stimulus-locked difference waveforms across blocks. Topoplots shows the mean difference wave at 500 ms (Cz).

ANOVAs revealed a significant region effect and a grammaticality effect in Block I, but there was no interaction between region and grammaticality. On the other hand, when the difference wave in Block II is observed, it can be easily seen that both grammatical and ungrammatical words were processed similarly. ANOVAs revealed that even though there was a significant region effect, no grammaticality effect was observed. Statistical results are summarized in Table 6.2 below.

Table 6.2: Stim-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

	Voltage (μV)			Statistics		
	Frontal	Central	Parietal	Gramm. df= 1,23	Region df= 2,46	Region × Gramm. df= 2,46
Block I	1.03 (.393)	0.95 (.294)	0.68 (.288)	F=11.436 p<.003 η^2=.332 1-β=0.875	F=31.415 p<.001 η^2=.577 1-β=0.999	F=0.573 p=.568
Block II	0.52 (.414)	0.27 (.295)	0.26 (.315)	F=1.253 p=.274	F=41.178 p<.001 η^2=.642 1-β=0.999	F=.720 p=.492

6.3.2.2 Response-locked waveforms

During the processing of a word, once participants had decided a stimulus was grammatical or ungrammatical, the next step was to press the button. When the waveforms are time-locked to the motor response (the button press), the brain response should reflect whatever processing mechanism is active from the time between the decision making and the motor response. This can be observed in the response-locked rare-minus-frequent difference waves in Figure 6.5 below. Firstly, a typical Bereitschafts Potential (BP) was observed specifically in the parietal region both in Block I and II, independent of the grammaticality effect. The BP, which is also called the readiness potential, reflects the activity in the brain prior to the behavioral motor response (the button presses) (Jahanshahi & Hallett, 2003). Next, in Block I, it can be seen that the difference wave peaks at -200 ms (prior to a motor response) which can be taken as an indication of differential processing of grammatical vs ungrammatical words.

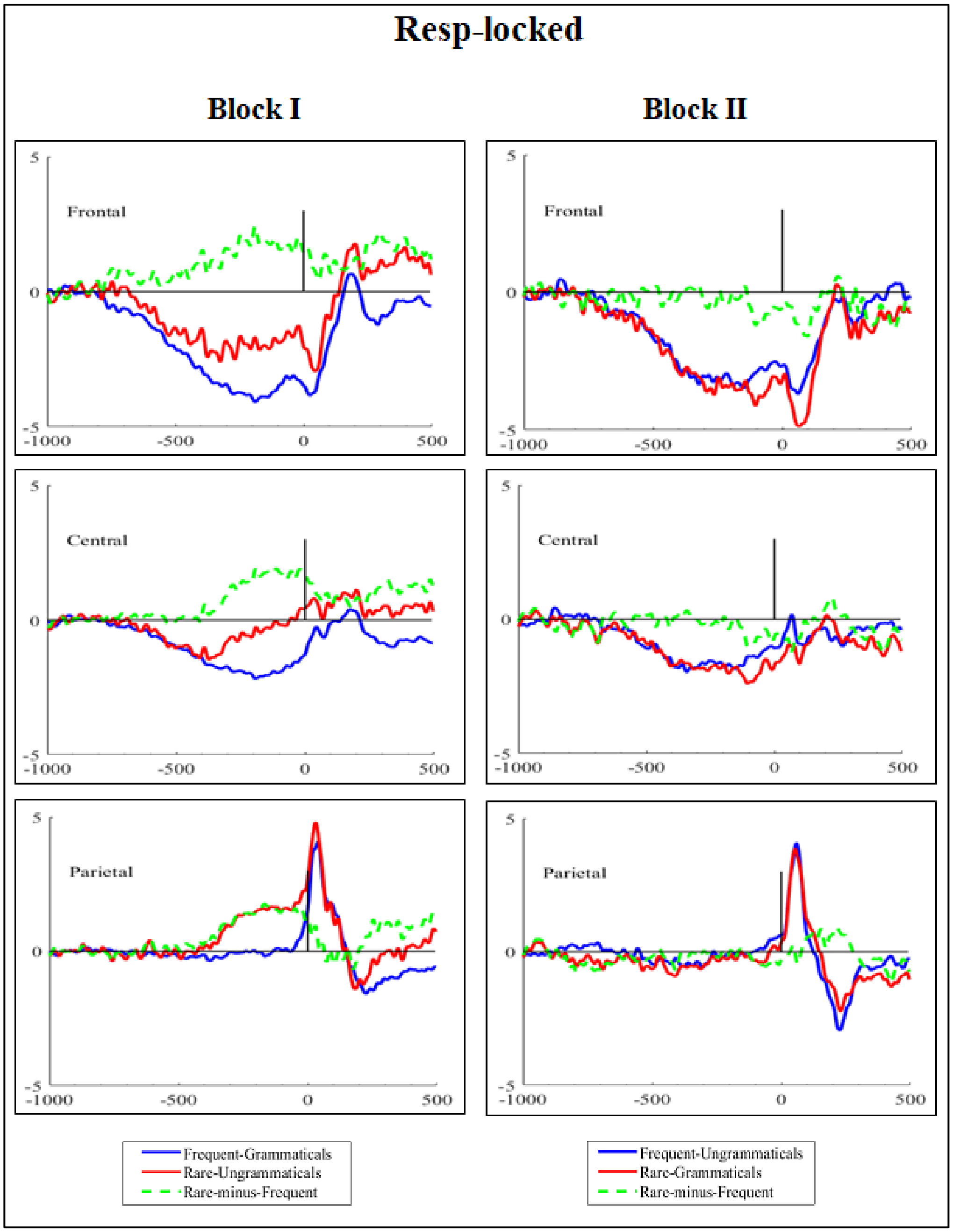


Figure 6.5: Response-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites across blocks. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV).

On the contrary, in Block II, there was no positive peak before the motor response and the difference wave again bounces around zero. This indicates that participants in Block II could not make a categorization decision before the button press. When the difference wave in Block II is observed, both grammatical and ungrammatical words were processed similarly. This must be due to the frequent presentation of ungrammatical words that might have affected the neural commitment. See Figure 6.6 below for the spatial distribution at -200 ms across the blocks.

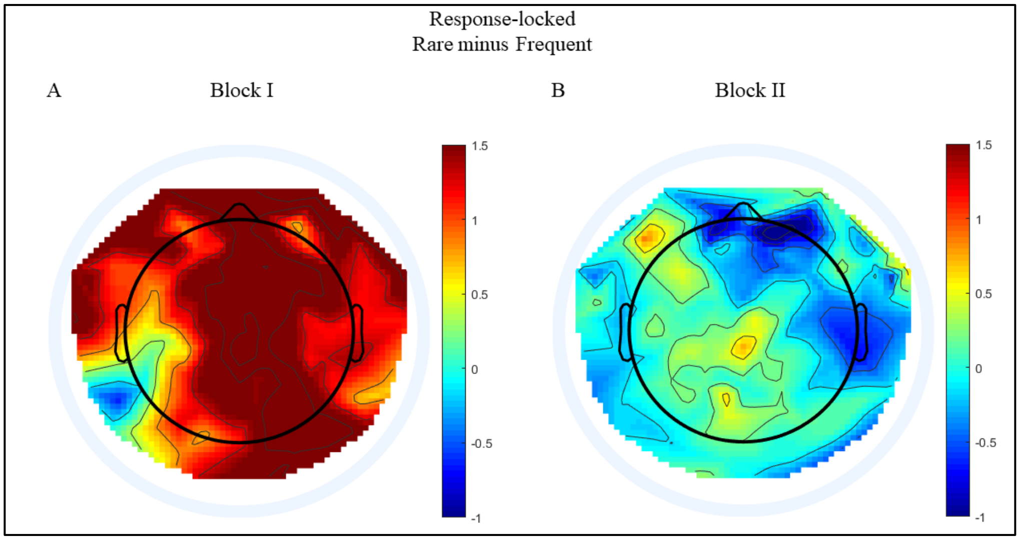


Figure 6.6: Spatial distribution of the response-locked difference waveforms across blocks. Topoplots shows the mean difference wave at -200 ms (Cz).

Statistical analyses mirror the stimulus-locked results in that there were a significant region effect and a grammaticality effect in Block I, but there was no interaction between region and grammaticality. As in the stimulus-locked averages, the amplitude of the difference wave was higher in the frontal region.

In Block II, the results revealed no grammaticality effect as opposed to a significant region effect. See Table 6.3 below for response-locked statistical results across blocks.

Table 6.3: Resp-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

	Voltage (µV)			Statistics		
	Frontal	Central	Parietal	Gramm. df= 1,23	Region df= 2,46	Region × Gramm. df= 2,46
Block I	1.93 (.607)	1.76 (.467)	1.62 (.598)	F=12.499 p=.002 η^2=.352 1-β=0.910	F=44.65 p<.001 η^2=.660 1-	F=0.240 p=.787
Block II	0.29 (.417)	0.26 (.341)	.09 (.310)	F=.525 p=.476	F=34.91 p<.001 η^2=.603 1-	F=0.191 p=.827

6.3.3 LPC Analysis

The distribution of the absolute waveforms after the offset of the word indicates the subsequent posterior positivity elicited by ungrammatical words in comparison to grammatical words. This observation paves the way for the second important result of this experiment: namely, the ungrammatical words show an extended positive deflection at the parietal region between 600 and 1000 ms (see Figure 6.7 below). The waveforms show an enhanced posterior positivity for ungrammatical words with a maximum at around 800 ms.

Statistical analyses revealed that in Block I, there is a significant difference between grammatical and ungrammatical words (t(23)=2.281, p=.032, d=.466, 1-β=0.715). This late positivity effect is due to the violation of sibilant harmony rule. Whereas, in Block II, no significant difference was observed (t(23)=.198, >.05). The absence of the LPC effect also showed neural commitment was somehow spoiled by the frequent presentation of ungrammatical stimuli.

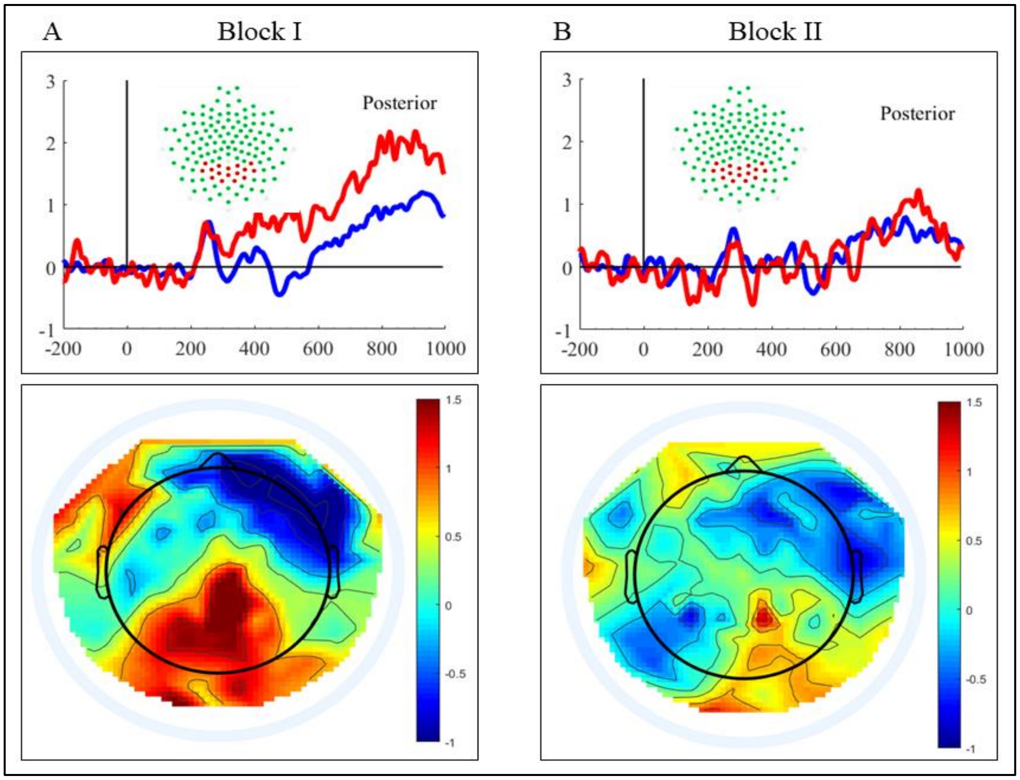


Figure 6.7: Higher panel: Grand average ERP absolute waveforms obtained for grammatical words (blue line) and ungrammatical words (red line) at the posterior region across blocks. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Lower panel: Spatial topography maps showing voltage differences between grammatical and ungrammatical words at 800 ms across two blocks. Left topo-plot indicates that the contrast between grammatical and ungrammatical words revealed a parietal positivity effect for non-words.

6.3.4 ERP-Behavioral Correlations

Correlations between P3 difference wave, LPC, and behavioral measures were computed across blocks. For the P3 wave, I correlated the mean amplitude of the rare-minus-frequent difference waveforms averaged across regions with mean reaction time, accuracy, and *d'*. For the LPC, I focused on the amplitude difference between the grammatical and ungrammatical words, averaged across the posterior region. For the behavioral measures, due to the fact that *d'* is already focused on the signal of deviant stimuli (ungrammatical words), I chose the accuracy and reaction of deviant stimuli to include in the correlation (see Luck et al., 2009, for further discussion). I used only the stimulus-locked ERP measures because the response-locked data deforms the ERP and reaction time relationship.

Table 6.4: Pearson r and associated p values for the correlations between behavioral measures, P3 and LPC ERP measures in Block I.

		d'	Acc	RT	P3	LPC
d'	Pearson's	-----	-----	-----	-----	-----
	p-value	-----	-----	-----	-----	-----
Acc	Pearson's	0.838***	-----	-----	-----	-----
	p-value	<.001	-----	-----	-----	-----
RT	Pearson's	0.410*	0.384	-----	-----	-----
	p-value	0.047	0.064	-----	-----	-----
P3	Pearson's	-0.180	-0.327	-0.450*	-----	-----
	p-value	0.401	0.118	0.027	-----	-----
LPC	Pearson's	0.088	-0.117	-0.180	0.593**	-----
	p-value	0.682	0.588	0.399	0.002	-----

* p < .05, ** p < .01, *** p < .001

In Block I (see Table 6.4 above), it was found that there was a reliable positive correlation between d-prime scores and accuracy and reaction time. This suggests that participants with stronger behavioral sensitivity showed higher accuracy but longer reaction times. Moreover, longer reaction times were associated with higher accuracy (marginally significant). There was also a negative correlation between reaction times and P3 amplitudes, implying that smaller P3 amplitudes were associated with longer reaction times. This indicates that the longer it takes for participants to evaluate the stimulus, the lesser the P3 amplitude – which depends on the number of attentional resources. This finding is consistent with the previous findings in that when the task difficulty is increased causing an increase in RT, P3 amplitude decreases (Polich, 2007).

Finally, there was a significant positive correlation between the P3 and LPC amplitudes which indicates that participants with stronger categorical perception in grammaticality showed a robust LPC, whereas those with less categorical perception showed a weak LPC. See Figure 6.8 below for a visual inspection of correlation matrixes of only the significant correlations.

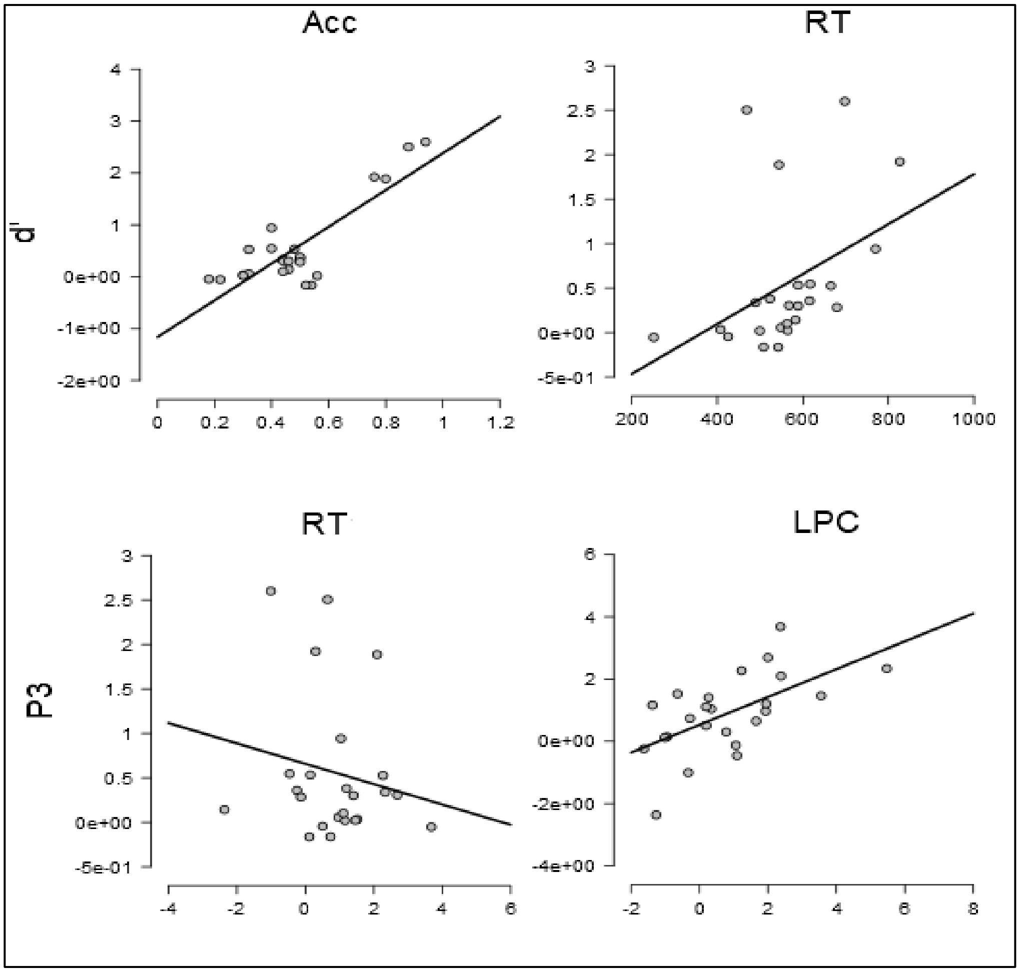


Figure 6.8: Correlation matrix for significant correlations in Block 1.

On the contrary, in Block II, since there was no significant P3 or LPC difference, the only significant correlation was between behavioral measures (d-prime and acc) which is a strong positive correlation (r=0.844, p<.001).

6.4 Discussion

This study used behavioral measures and ERP (P3 and LPC) difference waves to measure the underlying brain processes of phonological rule learning.

After a brief training session, participants were tested with a categorization task that required them to differentiate grammatical vs ungrammatical stimuli whose frequency was counterbalanced across two blocks. Behavioral results showed that participants learned the non-adjacent phonotactic rule after a brief training session. The ungrammatical words were detected with a sensitivity higher than zero, and there was a significant grammaticality effect on accuracy and reaction time across both blocks. This indicates that even though the frequency of the grammatical words changed across blocks, participants were still sensitive to the ungrammatical words implying a persistent rule learning effect. In both blocks, the bias was negative which indicates that participants were biased towards reporting that they had heard a grammatical word. It is important to note that this bias was not affected by the frequency of the stimuli. This indicates that participants were good learners even though their sensitivity was not that high. The behavioral results demonstrate that participants were able to learn this simple artificial phonotactic rule in the laboratory and that the acquired knowledge was measurable in their behavior.

In addition to behavioral measures, measuring the neurophysiological correlates of this learning gives us insight into the brain processes that underlie phonotactic pattern learning. What these ERP results shows is whether the newly acquired knowledge is directly “channeled” into incremental, real-time phonological predictive processing at the neural level or not. This can shed light on the time course of the processes that lead to a behavioral response. For example, for ungrammatical stimuli, is a violation detected immediately at the parsing of a violating segment, or much later?

EEG results showed a clear AEP to the onset of the sound and a Bereitschafts Potential prior to the onset of the button-press in both blocks. However, while changing the frequency of the grammatical and ungrammatical words did not affect behavioral discrimination, it affected the neural responses. While grammatical words were frequent (Block I), rule learning was reflected in both a P3 difference waveform and an LPC, whereas when ungrammatical words were presented frequently, this difference at the neural level was lost. In the following paragraphs, I will discuss the presence of the neural responses in Block I and then try to discuss the reasons for the absence of a response in Block II.

Throughout the test phase, the violating segment within the ungrammatical stimulus appeared at 200 ms. In Block I, after the offset of the word (at 400ms), grammatical and ungrammatical words started to be processed differently. While both the grammatical and ungrammatical words' waveforms were negative-going, the rare grammatical words were more positive. Once the rare-minus-frequent difference waveform was created, it can be seen that the peak of the P3 difference wave appeared at 500 ms. This peak indicates the point in time when the participant has categorized the stimulus as grammatical or ungrammatical, prior to the behavioral response (one needs to determine grammaticality first to choose a response selection). This implies that the subject's brain understands that a particular stimulus is violating the pattern exactly at the violation point, not after the whole word has been processed. In other words, the phonological pattern detection mechanism does not wait until the end of the word to rule out an ungrammatical word. Thus, this is also evidence that phonological parsing is incremental. This time course information cannot be inferred from behavioral measures alone. Furthermore, when time-locked to the subject's response (the button presses), a positive peak immediately preceded the response. This wave reflects the decision activity between the categorization and the button press response, which consists of selecting, preparing, and executing the appropriate response. Interestingly, none of these neural processes were observed in Block II, where the participants showed categorization differences behaviorally. This indicates that once the frequency of the grammatical and ungrammatical words changed, somehow the categorization mechanism started to work differently. Thus, our question of whether the frequency of stimuli presentation affects neural commitment can be answered: After a brief training, when participants were bombarded with more ungrammatical stimuli, the neural commitment was spoiled but the rule knowledge at the behavioral level remained intact. While the neural encoding of the first categorization mechanism was accessible to EEG in Block I, it was not accessible to EEG in Block II. These results imply that the neural encoding of the categorization mechanism across blocks was different. Possibly, after the training session, when participants were tested with more grammatical stimuli, they were depending more on linguistic rule knowledge; and when they started to get more ungrammatical stimuli, they relied on a different categorization mechanism. Behaviorally both of these categorization mechanisms work similar, but at the neural level, they work differently; while the first is accessible to EEG (measured by P3 difference waveform), the latter is not.

In Block I, LPC modulations indicate that phonotactic constraint violations affected later stages of cognitive processing. This is really important because the LPC is elicited in response to phonotactic violations even when ungrammatical words have already been categorized as violating the pattern. Our P3 difference waveform just showed the categorization difference between the two types of stimuli. However, LPC elicitation adds something more valuable on top of this which is that the LPC is reflecting anomaly detection in rule-governed phonological sequences. Lei, Wang, Zhu, Chen, & Li (2019) also found a P3 and a late positive-going slow-wave (PSW) for category-based induction which was implied to reflect the distance effect (non-adjacent dependency in our context) in inductive reasoning. This distance effect can be thought of as the distance between the two sibilants in this experimental context. Another implication of LPC elicitation in the current experimental context is that during the training phase, participants must somehow grammaticalize the phonotactic aspects of the words. This can be taken as an indication of the instantiation of rule-based knowledge into the learner’s real-time language processing system (Domahs et al., 2009). Once the phonotactic forms have been grammaticalized in the training or during the consecutive presentation of grammatical words in Block I, violations of these forms elicited a P600-like an effect. This grammaticalization idea also explains the absence of an LPC in Block II where the consecutive presentation of ungrammatical words must have deformed the grammatical forms that had been constructed during training. That may have caused the absence of P600-like effect.

Eliciting a P3 requires an oddball (rare-vs-frequent) experimental design. In Block I, the ungrammatical words appeared as infrequent ‘deviants’, so the grammatical-ungrammatical categorization coincides with a frequent-rare categorization. However, the difference wave obtained (grammatical vs ungrammatical) is not simply due to tracking the relative frequency of the stimuli. The difference reflects an underlying sensitivity to the phonotactic rule. Although the grammatical words were presented at a higher frequency than the ungrammatical, the relative frequencies cannot be informative to participants unless they understand the rule. Participants cannot track the relative frequencies without first recognizing the difference between grammatical and ungrammatical stimuli. Without having learned the rule, there is no phonetic or phonological feature that separates frequent grammatical tokens from the infrequent ungrammatical tokens.

Another point that needs clarification is why the absolute P3 waves were negative and why the spatial distribution of this effect was frontal-central. Firstly, only exogenous ERPs have an interpretation of absolute waveforms (brain stem response, like the AEP): the P3 is an ERP that reflects a cognitive mechanism. Secondly, studies may vary depending on the modality of the presentation, the design, or the input parameters (see Friederici & Meyer, 2004), thus, the expectation that an effect should occur in a specific polarity is problematic. A frontal-central distribution of the P3 effect due to the modality of stimulus presentation or availability of the attentional resources is discussed in Polich (2007) as well. For example, for a different ERP component (N400), Holcomb & Neville (1990) reported a broader N400 distribution in auditory as compared to visual presentation.

To sum up, when the sibilant harmony rule was simplified (CV.CV words) and the design was changed to incorporate categorical learning, I measured the neural commitment of this phonological rule learning via P3 and LPC. The behavioral results indicate that the participants learned the rule. The neural results indicate that this learning took the form of a neural commitment. Participants learned the rule and used it to make active predictions by categorizing words as ungrammatical around the point of violation. This ability must be instantiated at the neural level, meaning rapid neural tuning has occurred in this lab setting. Changing the frequency affected the nature of this neural commitment while not affecting the behavioral results. However, the question remains of where this neural commitment takes place. The fact that non-adjacent rules of this type have not been found to elicit MMN effects suggests that this neural commitment is not made at the perceptual level (in the auditory cortex) (of course, only for the longer words). Instead, these more complex phonological patterns may be neurally encoded at a 'higher' level. Because the P3 is an attention-driven comparison process, the rule must be accessible to working memory, although it may not be accessible to the auditory sensory memory.

6.5 Conclusion

The aim of the current study was to correlate phonological rule-learning with a neurophysiological response by using a simplified version of the SH rule. Previous research has found several types of evoked responses to violations of lab-learned phonological patterns, including the MMN, the LPC, and the P3.

The particular response observed in each case is dependent on several factors, including study design, the mode of presentation, the type of stimulus, and the pattern or rule presented. I exposed participants to a non-adjacent phonotactic pattern in the form of CV.CV words in an oddball paradigm, and found the evidence of learning in the participants’ behavioral responses (categorization above chance level), as well as a robust brain response reflected in a P3 difference waveform and LPC. This study was a preliminary design that showed a neurophysiological measure for further investigation of phonological rule-learning. Future work will also be necessary to determine exactly what kinds of rules and rule-learning can be measured in this way.

Chapter 7

EXPERIMENT 5: NEURAL MARKERS OF EXPLICIT RULE LEARNING

In this chapter, the results of the fifth experiment that was conducted to observe neurophysiological markers of phonological rule learning will be presented. After the experiment with simpler stimuli (CV.CV) showed both a P3 as a marker of categorization difference and an LPC indexing phonological anomaly detection, we decided to replicate these results with a different learning mechanism – namely with a different training condition that involved explicit instruction. The original idea was to see whether we observe a more robust ERP by a strong behavioral rule knowledge. This strong behavioral rule knowledge will be ensured by explicit instruction. In this respect, the aim of this experiment is to look for the neurophysiological markers of an explicit phonological rule learning mechanism. Investigating the differences between implicit (Experiment 4) and explicit phonotactic learning (Experiment 5) is important for finding answers to the following questions: Are different learning processes encoded differently at the neural level? If so, how does this relate to the acquisition of a first language or second language?

It has been established that learning a second language is much more difficult than learning the first language. First language learning is implicit in the sense that it develops without depending on central attentional resources (conscious attention); whereas second language learning is explicit, meaning that it makes heavy demands on working memory (Martin & Ellis, 2012). Moreton & Pertsova (2016) differentiate these two learning mechanisms by calling them the explicit system and the implicit system. According to their study, the implicit learning system is “effortless, unconscious, gradual, and cue-based; it does not need focused attention or working memory, and its use is facilitated by training without feedback, instructions that do not mention rules, and non-verbalizable stimulus features”. Whereas the explicit learning system is “effortful, conscious, abrupt, and rule-based; it demands focused attention and working memory, and its use is facilitated by training with right/wrong feedback, instructions to seek a rule, and the use of easily verbalizable stimulus features” (Moreton & Pertsova (2016: p. 277)). Moreton et al. (2017) later associated these two systems with cue-based vs rule-based learning models. In the cue-based model, the learner uses cues scattered in the data to make an inference, contrary to the rule-based model where the learner decides directly by using a rule. Pater (2014) refers to this distinction by relating phonotactic learning to implicit learning because the learner is trained with only positive examples. Therefore, phonotactic learning is cue-based. Pater notes that this distinction also refers to procedural (implicit, cue-based) vs declarative memory systems (explicit, rule-based) (Pater, 2014). It is evident that these two learning processes have different properties that are triggered by different conditions (Maddox & Ashby, 2004). However, less is known about how these two learning systems are encoded at the neural level: Is there a difference in neural coding between implicit and explicit learning?

The distinction between implicit and explicit systems also concerns the different learning mechanisms which are pertinent to the dichotomy of domain-specific and domain-general learning systems. According to Campbell & Tyler

(2018), language-specific functions such as syntax (Zaccarella & Friederici, 2015) are carried out by specific regions whereas cognitive functions such as attention and memory are carried out by other domain-general neural networks. My hypothesis is that studying the neural commitment of explicit learning of phonotactic patterns will pave the way for understanding the neural encoding of the two different learning mechanisms. In this respect, this chapter's main question is as follows: Do implicit and explicit learning lead to different neural commitments?

Contrary to the implicit learning in Experiment 4 where participants merely repeated grammatical exemplars without any explicit instruction, participants in the current experiment, namely the explicit group, had the rule explained. The experimental details are explained in section 7.1. The results, presented in section 7.2, showed that while the explicit group performed much better behaviorally, they did not exhibit an ERP response modulated by well-formedness. The comparative results between Experiments 4 and 5 are also presented. I discuss the implications of these results in section 7.3, and section 7.4 concludes the chapter by reporting that explicit learning did not engage neurophysiological mechanisms that lead to prediction models at the neural level. This suggests that implicit lab-learning experiments tap into the kind of unconscious, automatic learning that is characteristic of natural language acquisition.

7.1 Methods

7.1.1 Participants

A total of 23 University of Delaware students were recruited as participants and provided written consent for participation in the experiment. 2 participants were eliminated due to exceptionally noisy EEG data, yielding a final sample of 21 participants. Each participant received course credit for participation. 18 of the 21 participants were female. Four participants were left-handed. The mean age was 20 (*SD=1.22, range=18 to 22*). None of the participants reported a history of hearing loss or speech/language impairments, and all reported speaking English as their first and only language.

7.1.2 Stimuli

The stimuli were the same as in Experiment 4, namely a simplified version of the Sibilant Harmony pattern in the form of CV.CV words.

7.1.3 Apparatus and procedure

The apparatus was the same as in the previous experiments. The procedure for the experiment consisted of three phases: a familiarization phase, a training phase, and a testing phase. In the familiarization phase, firstly, participants were told that they will learn a new language and the language they will learn have a rule which is “s and ʃ cannot appear in the same word”. And then, they entered a familiarization session in which they had presented all the harmonic and disharmonic words and were instructed to press a button in response to each stimulus to categorize the stimulus. After they pressed the button to categorize each word, the feedback was given informing them that they had pressed the correct or incorrect button. Once this phase ended, they entered the training and test sessions which were exactly the same as Experiment 3. The only difference was that participants received only one block in which a deviant (ungrammatical) stimulus appeared infrequently among repeated occurrences of a standard (grammatical) stimulus. Participants discriminated between grammatical and ungrammatical words, with grammatical words appearing in 83% of trials (1000 words) and ungrammatical words appearing in 17% of trials (200 words), with 1200 total trials. The duration of the test session was approximately 40 minutes.

7.1.4 Data recording and analysis

Behavioral and electrophysiological data recording and analysis were exactly the same as in Experiment 4.

7.2 Results

7.2.1 Behavioral Results

Ungrammatical words were detected with a mean sensitivity of 1.613 (d') (*SD=1.20*), a score significantly greater than zero, t(20)=6.12, p<.001, d=1.336, 1-β=0.999. Mean bias rate was -0.311 (c) (*SD=.29*) – negative and significantly different than zero which is expected and the artifact of the oddball paradigm. See Figure 7.1 below for a visual comparison. This result indicates that all the deviants were detected with a mean sensitivity higher than zero. Sensitivity for ungrammatical words was higher than zero, showing that participants learned the rule. Mean accuracy (percent correct) for grammatical words was .83 (*SD=.13*) and .65 (*SD=0.22*) for ungrammatical. The difference (based on mean acc) was significant t(20)=4.808, p<.001, d=1.049, 1-β=.998. The means accuracy of both word types was higher than chance (0.5). Mean RT was 549 ms (*SD=98*) for grammatical words and 619 ms (*SD=108*) for ungrammatical. The difference was significant, t(20)=7.494, p<.001, d=1.635, 1-β=1.0.

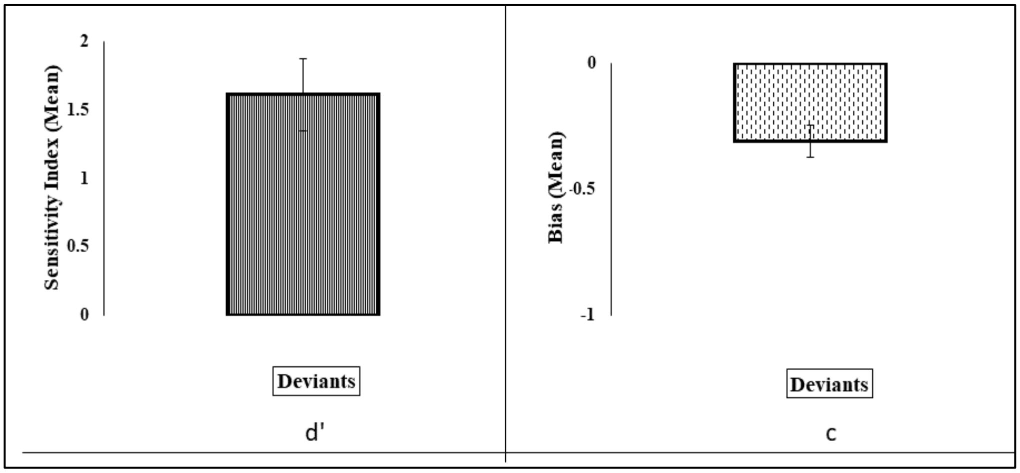


Figure 7.1: Averages of sensitivity index rates (left panel) and bias rates (right panel). Error bars indicate standard error of the mean.

The behavioral results above were compared to the results of implicit learners in Experiment 4. The sensitivity index and bias rate means were compared in an independent-samples t-test. The results showed that there was a significant difference between the two groups of learners, t(43)=3.47, p=0.001, d=1.038, 1-β=0.924; in that the explicit group had a better detection ability. However, the mean bias rate was not significantly different between the two groups, t(43)=0.217, p>0.05. See Figure 7.2 below for a visual comparison. One point that that needs to be clarified is that Experiment 4 had 300 trials but Experiment 5 had 1200 trials. Thus, in order to make a direct comparison between the two experiments, we need to look at the first 300 trials in Experiment 5. When the behavioral results of the first 300 trials were analyzed, it has been concluded that explicit learners had the same level of sensitivity from the very beginning to the end[7].

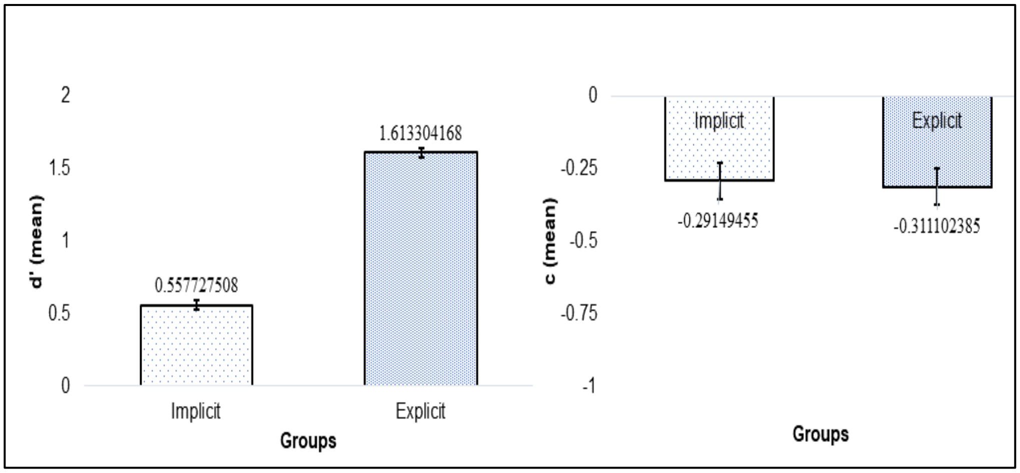


Figure 7.2: Sensitivity and bias rate means comparison between implicit learners vs explicit learners.

7.2.2 P3 Rare-minus-frequent difference waveforms

7.2.2.1 Stimulus-locked waveforms

Figure 7.3 below shows stimulus-locked ERP absolute waveforms and the difference waves (rare-minus-frequent). It can be seen that both frequent and rare stimuli revealed a typical auditory evoked potential (AEP) at frontal-central

---

[7] This was tested in a one-way repeated measures ANOVA using the trials as an independent variable that had four levels: (i) 1-300, (ii) 301-600, (iii) 601-900, (iv) 901-1200. Results showed that there was no effect of trial on sensitivity index rates, F(3,60)=0.235, p=0.872.

electrodes. It was important to get this AEP to the onset of the word, whether it was grammatical or ungrammatical because it indicates that the auditory system has responded to the sound coming out of silence.

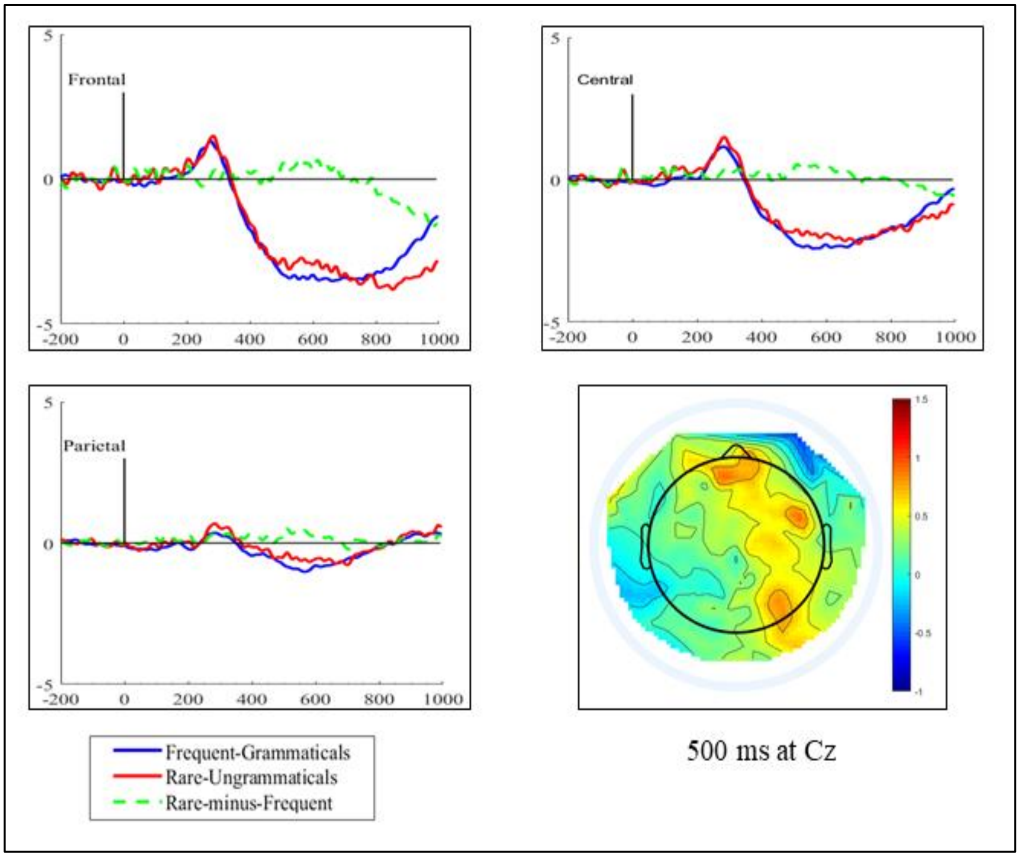


Figure 7.3: Stimulus-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Topoplot shows the mean difference wave at 500 ms at Cz.

Figure 7.3 also shows that there is a small difference between grammatical (blue) and ungrammatical (red) words beginning from the offset of the word (400ms) to 700 ms specifically at the frontal and central regions. However, this difference was not statistically significant. A whole scalp snapshot of the difference, which does not show any frontal-central positivity, can also be viewed in the same figure above. ANOVAs revealed a significant region effect but no grammaticality effect, and there was no interaction. Without a significant grammaticality effect, the region effect is not interpretable. Statistical results are summarized in Table 7.1 below.

Table 7.1: Stim-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

Voltage (μV)			Statistics		
Frontal	Central	Parietal	Gramm. df= 1,20	Region df= 2,40	Region × Gramm. df= 2,40
.329	.288	.220	F=3.480	F=27.646	F=0.279
(.187)	(.156)	(.172)	p=.077	p<.001	p=.758
			η^2=.148	η^2=.580	
				1-β=0.990	

Moreover, in order to compare the results of Experiment 4 with Experiment 5, a 2 x 2 ANOVA with the group as an independent factor (implicit vs explicit) and grammaticality of the word (grammatical vs ungrammatical) as a within-subjects factor was run. The analysis included the stimulus-locked voltage values only from the central region to observe the P3 effect. Results revealed a main effect of grammaticality (F(1,43)=12.675, p<.001) in the predicted direction, namely ungrammatical words caused more positivity in the implicit group. A marginally significant interaction between group and grammaticality was also revealed (F(1,43)=3.609, p<.064). However, there was no main effect of group as a between-subject factor, (F(1,43)=.986, p<.326). These results show that grammatical-ungrammatical difference was clear in the implicit group but not in the explicit group. This comparison was made over the 1200 trials of the explicit experiment vs 300 trials of the implicit experiment. A visual comparison of the P3 effect under the same number of trials (300 vs 300) showed the same results, see Figure 7.4 below.

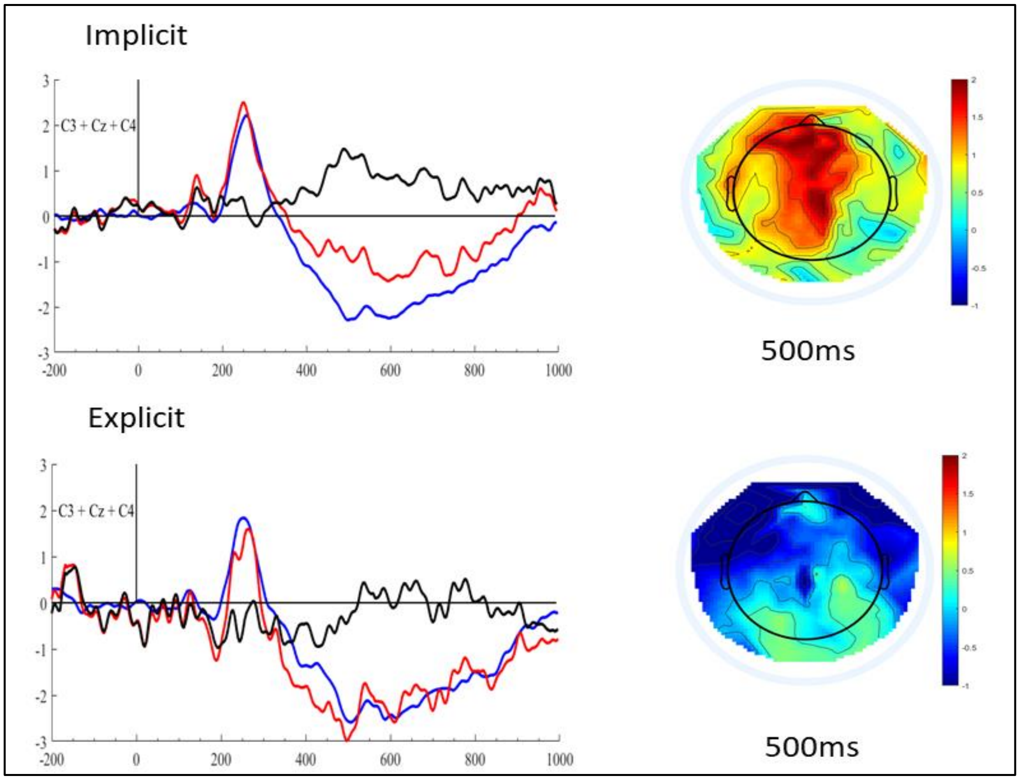


Figure 7.4: Visual comparison of the P3 effect between the implicit vs explicit groups with same number of trials (300 trials). Grand average ERP absolute waveforms obtained for grammatical words (blue line), ungrammatical words (red line), and the difference (black line) at the central region. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Spatial topography maps showing voltage differences between grammatical and ungrammatical words at 500 ms.

7.2.2.2 Response-locked waveforms

Figure 7.5 below shows response-locked ERP absolute waveforms and the difference waves (rare-minus-frequent).

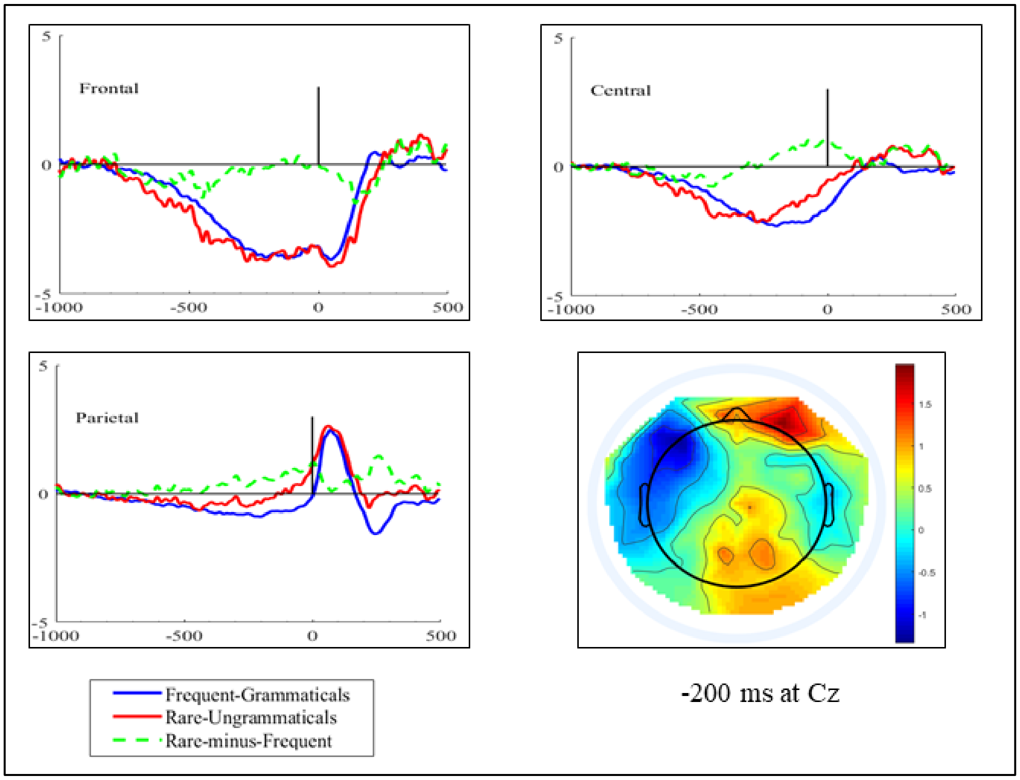


Figure 7.5: Response-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Topoplot shows the mean difference wave at -200 ms at Cz.

A typical Bereitschafts Potential (BP) was observed specifically in the parietal region, independently of the grammaticality effect. Next, it can be seen that in the central region the difference wave was positive and peaked around -200 ms which can be taken as an indication of differential processing of grammatical vs ungrammatical words. At the frontal and parietal regions, there was no positive peak before the motor response, and the difference wave was not significantly different from zero. When the difference wave is observed, it can be seen that both the grammatical and ungrammatical words were processed similarly. The spatial distribution at -200 ms at Cz also showed no positivity difference. Statistical analyses mirror the stimulus-locked results in that there was a significant region effect, but no grammaticality effect or interaction. See Table 7.2 below for response-locked statistical results.

Table 7.2: Resp-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

Voltage (μV)			Statistics		
Frontal	Central	Parietal	Gramm. df= 1,20	Region df= 2,40	Region × Gramm. df= 2,40
.023	.632	.629	F=1.317	F=28.844	F=1.838
(.575)	(.397)	(.252)	p=.265	p<.001	p=.172
			η^2=.062	η^2=.591	
				1-β=0.992	

In addition, in order to compare the results of Experiment 4 with Experiment 5, a 2 x 2 ANOVA with the group as an independent factor (implicit vs explicit) and grammaticality of the word (grammatical vs ungrammatical) as a within-subjects factor was run. The analysis included the response-locked voltage values only from the central region to observe the P3 effect. Results revealed a main effect of grammaticality (F(1,43)=14.818, p<.001) in the predicted direction, namely ungrammatical words caused more positivity in the implicit group. A marginally significant interaction between group and grammaticality was also revealed (F(1,43)=3.304, p<.076). However, there was no main effect of group as a between-subject factor, (F(1,43)=2.046, p<.160). These results show that grammatical-ungrammatical difference was clear in the implicit group but not in the explicit group. This comparison was made over the 1200 trials of the explicit experiment vs 300 trials of the implicit experiment. A visual comparison of the P3 effect under the same number of trials (300 vs 300) showed the same results.

7.2.3 LPC Analysis

As in the case of the P3 difference waveform, the ungrammatical words showed an extended positive deflection in the parietal region between 600 and 1000 ms. See Figure 7.6 below. However, no statistically significant difference was observed between the grammatical and ungrammatical words, t(20)=1.263, p=.221, d=.276. The explicit rule learners detected the anomaly in response to the violation of the sibilant harmony rule at the behavioral level. However, no significant brain response was measured.

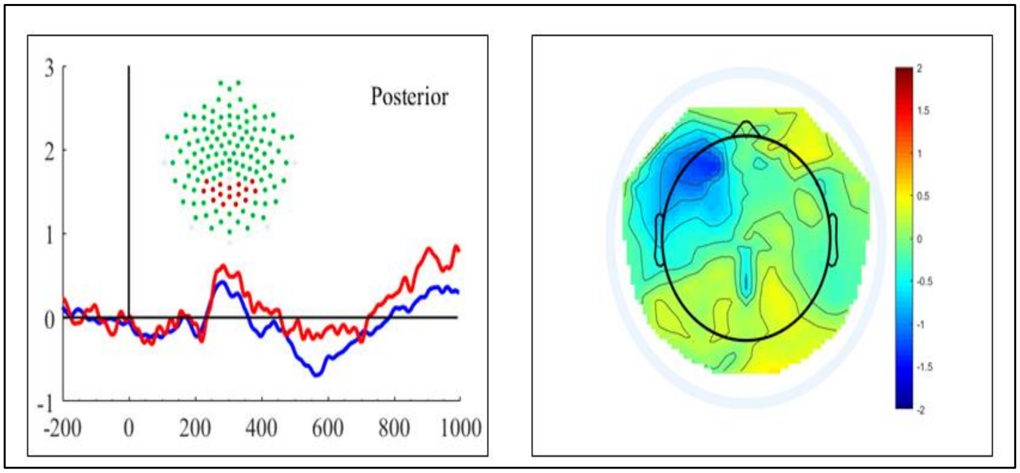


Figure 7.6: Left Panel: Grand average ERP absolute waveforms obtained for grammatical words (blue line) and ungrammatical words (red line) at the posterior region. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Right panel: Spatial topography maps showing voltage differences between grammatical and ungrammatical words at 800 ms.

Topo-plot indicates that the contrast between grammatical and ungrammatical words did not reveal significant parietal positivity effect for non-words. Furthermore, in order to compare the results of Experiment 4 with Experiment 5, a 2 x 2 ANOVA with the group as an independent factor (implicit vs explicit) and grammaticality of the word (grammatical vs ungrammatical) as a within-subjects factor was run. The analysis included the stimulus-locked voltage values from the posterior region to observe the LPC effect. Results revealed a main effect of grammaticality (F(1,43)=6.343, p=.016) in the predicted direction, namely ungrammatical words caused more positivity in the implicit group. There was also a main effect of group as a between-subject factor, (F(1,43)=6.409, p=.015). However, there was no interaction between group and grammaticality, (F(1,43)=1.487, p=.229). These results show that grammatical-ungrammatical difference was clear in the implicit group but not in the explicit group. This comparison was made over the 1200 trials of the explicit experiment vs 300 trials of the implicit experiment. A visual comparison of the LPC effect under the same number of trials (300 vs 300) showed the same results, see Figure 7.7 below.

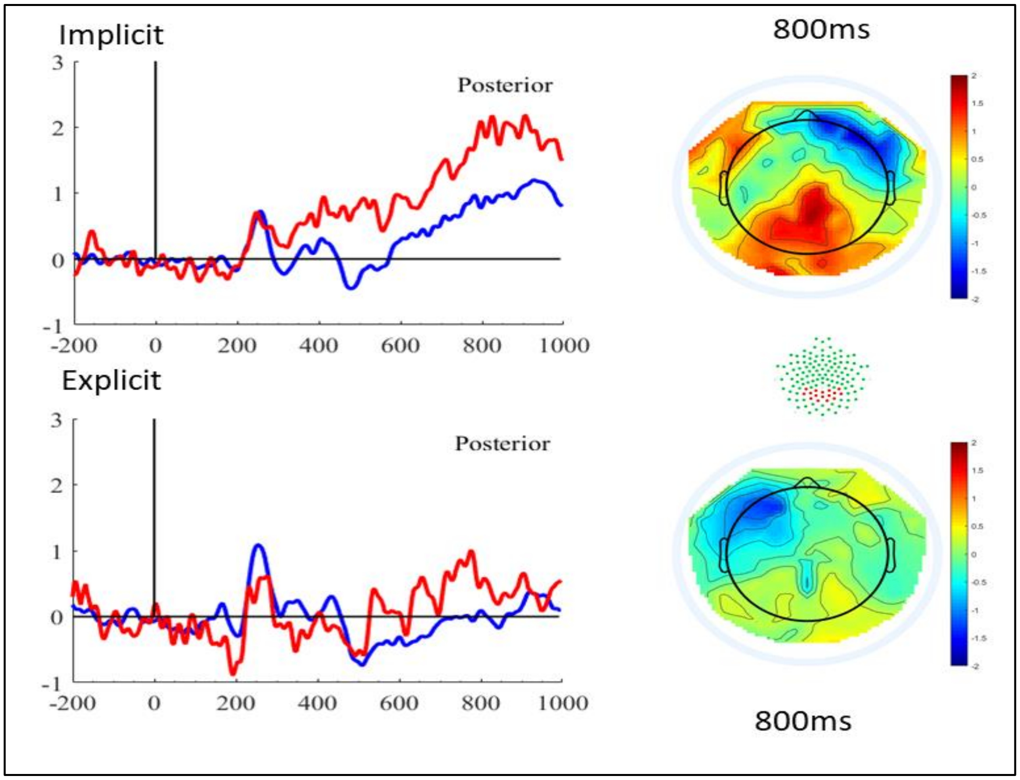


Figure 7.7: Visual comparison of the LPC effect between two the implicit vs explicit groups with same number of trials (300 trials). Grand average ERP absolute waveforms obtained for grammatical words (blue line) and ungrammatical words (red line) at the posterior region. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in μV). Spatial topography maps showing voltage differences between grammatical and ungrammatical words at 800 ms.

7.3 Discussion

The aim of this experiment was to look for the neurophysiological marker of the explicit phonological rule learning mechanism. In this study, we tried to answer the following questions: Are different learning processes encoded differently at the neural level? If so, how does this relate to the acquisition of a first language or second language? The hypothesis behind these research questions was that when the neural commitment of explicit learning of phonotactic patterns is analyzed and compared to that of implicit learning, the findings will pave the way for understanding the neural encoding of the two different learning mechanisms. Specifically, the main question that is being asked in this chapter regards the different neural commitments caused by implicit and explicit learning.

The main difference between this explicit experiment from the previous implicit experiment is that participants were told the rule before the training session. Therefore, during training, they were supposed to use a different learning mechanism compared to that of the implicit group. Behavioral results showed that participants' sensitivity to the ungrammatical words was very high – specifically, the explicit group was much more sensitive to the violation of the phonotactic rule than the implicit group. The bias measure was negative which indicated that participants were biased towards reporting that they had heard a grammatical word. As for accuracy, there was a significant difference between the mean rating of grammatical words vs ungrammatical words; participants were more accurate when they were categorizing the grammatical words, as expected. And finally, reaction time results followed the accuracy results in that it took participants longer to categorize the ungrammatical words.

These behavioral results demonstrate that as in second language learning, when learners are told the rule first, they learn better. When the behavioral results of the explicit learners in Experiment 5, were compared to that of the implicit learners in Experiment 4, it can be concluded that after being told the rule, the learning mechanism that the explicit group incorporated during the training session was better than the learning mechanism used by the implicit group. When we go back to the distinction between cue-based (implicit) vs rule-based (explicit) systems (Moreton et al., 2017), following these results we can say that at the behavioral level, the rule-based learner – the learner that decides directly by using the rule – is a better learner. Furthermore, since the explicit learner demands more focused attention and working memory, this learning is carried out by domain-general neural networks. Thus, at the behavioral level, domain-general learning is much better.

On the contrary, the neural results showed the reverse. The explicit learners showed no P3 modulation, despite the presence of a robust AEP in stimulus-locked waveforms and Bereitschafts Potential in response-locked waveforms. Both the stimulus-locked and the response-locked P3 difference waveforms showed that grammatical and ungrammatical words were processed similarly, yielding no significant grammaticality effect, despite the robust region effect. However, without a significant neural categorization difference between the two types of words, this region effect was uninterpretable. The lack of ERP differences could not be due to the fact that the ungrammatical words got repeated so many times in the experiment and became perceptually similar to the grammatical words because participants consciously knew which category they belong to, as it can be seen from the behavioral results.

As was pointed out in Chapter 6, LPC modulations indicate that phonotactic constraint violations affected later stages of cognitive processing. This is important because while the P3 difference waveform showed the categorization difference, the LPC is elicited in response to phonotactic violations even when ungrammatical words have already been categorized as violating the pattern. Therefore, the LPC elicitation adds something more valuable on top of the P3 elicitation (anomaly detection mechanism is still active even after the word was categorized as ungrammatical). Also, it has been demonstrated in the literature that native speakers', L2 learners', and lab learners' processing of phonological violations elicit the LPC (McLaughlin et al. (2010); Domahs et al. (2009); Moore-Cantwell et al. (submitted)). However, explicit learners did not show any LPC difference. The absence of an LPC here implies that neural substrates of explicit rule learning cannot be measured with EEG. This implies that explicit learners did not grammaticalize the phonotactic aspects of the words.

In order to directly compare the neural commitment of implicit learners to that of explicit learners, we examined the neural commitment of the first 300 trials in Experiment 5. These neural results also showed no significant P3 or LPC component in response to the rule violation during the first 300 trials of Experiment 5. These results are interpreted to indicate that implicit and explicit learning leads to different types of neural encoding of the acquired phonotactic rule. Based on these results it can be deduced that implicit learning leads to a measurable neural learning response typical of categorization systems, whereas explicit learning leaves the brain silent. Of course, the brain must be doing something because there was behavioral evidence; however, this explicit learning mechanism is related to some higher-level, domain-general mechanism, which was supported by a different brain region that may not be accessible to EEG, like the basal ganglia. This interpretation is in line with Moreton et al. (2017)'s distinction between cue-based (implicit) and rule-based (explicit) models; the former is more like typical phonotactic learning whereas the latter is classic visual category learning which depends on frontal-striatal circuits (Ashby & Maddox, 2005) that may not be reflected on an EEG. On the contrary, the study by Yang & Li (2012) showed activation in the frontal-striatal network for implicit learning. The study tested the neural correlates of explicit and implicit learning of artificial grammar sequences by an fMRI technique. It has been revealed that both learning processes showed activation in cortical and subcortical areas; while explicit learning showed activation in a network that uses the insula, implicit learning caused activation in the frontal-striatal network. The paper also discussed the role of individual differences in working memory as a key factor affecting the two types of learning processes. All in all, I conclude that deeper regions of the brain cannot be measured with EEG, since it only picks up activity from the cortex.

Another point that needs to be discussed is the nature of neural commitment. My claim is that although implicit learning leads to moderate behavioral sensitivity, it creates a strong neural commitment which was accessible to EEG and probably lasts longer. On the contrary, while the explicit learning leads to strong behavioral sensitivity, it creates a weak neural commitment which was not accessible to EEG and may not last as long. It would be really informative to test the explicit learner's behavioral sensitivity one week after the training. My assumption is that after one week the implicit group's sensitivity will stay the same but explicit group's sensitivity will definitely decrease due to the weak neural commitment. Another possibility is that since participants were familiarized with the test task with feedback, they might just memorize the correct answers and have a different strategy when being tested later. This different strategy may be the cause of different neural commitment.

To sum up, these results fit with the observation that first language acquisition is implicit and leads to long-term neural encoding, whereas adult 2^{nd} language acquisition, based on explicit learning, leads to a fundamentally different knowledge state. I conclude that lab-based learning experiments mimic naturalistic long-term implicit language learning.

7.4 Conclusion

The aim of this chapter was to examine the neurophysiological correlates of explicit learning of a non-adjacent phonotactic pattern. The explicit rule learning mechanism was tested with explicit instruction – specifically, the learners had the rule explained. While recording EEG, participants were presented with words that were either well-formed or ill-formed according to the rule. It was found that explicit learners performed behaviorally with accuracy levels indicating knowledge of the rule. However, they did not show an ERP response modulated by well-formedness, which would be interpreted as reflecting categorization decision and anomaly detection. These results, together with the results from Chapter 5, show that implicit and explicit learning converges on similar knowledge states, but with different underlying neural mechanisms. This suggests implicit lab-learning experiments tap into the unconscious, automatic learning that is characteristic of natural language acquisition.

Chapter 8

EXPERIMENT 6: NEURAL MARKERS OF COMPUTATIONALLY COMPLEX PATTERNS

In this chapter, I will present the results of the sixth experiment that was conducted to test the computational complexity effect on the neural system. The sibilant harmony pattern in the form of *CV.CV* words (i.e., sasa or ʃaʃa) is both (strictly) local and (strictly) non-local (piecewise) in terms of substrings and subsequences respectively. When the strings are analyzed with a window of three elements (k-factors), the pattern would be a local pattern with k=3, namely SL3. For example, the 3-factors of *sasa* are {#sa, sas, asa, sa#}. If these are allowed than *sasa* is a well-formed word. At the same time, when the sequences (non-local elements) are analyzed with a window of two (k-factors), the pattern would be a strictly piecewise pattern with k=2, namely SP2. For example, the 2-factors of sasa are {s…a, a…s, s…s, a…a}, note that this time the elements are non-adjacent. If these are allowed, then sasa is a well-formed word. Another way of looking at the *CV.CV* word forms in terms of whether they are SL3 or SP2 is defining the constraints. When we define a constraint like *sVʃ or *ʃVs, meaning that "do not have s followed by ʃ with an intervening vowel", this would be a definition of a local relation. On the contrary, when we define a constraint like *s…ʃ or *ʃ…s, which means that "do not have s followed by ʃ (the distance between the segments does not matter at all)", this would be a definition of a non-local (piecewise) relation.

In Experiment 4, I showed that, in response to words in the form of CV.CV that violates sibilant harmony, the P3 difference waveform indexes categorization between harmonic and disharmonic words, and the LPC indexes anomaly detection in a rule-governed phonological sequence. However, a key question in this context will be raised as follows: Are the P3 and LPC indexing violations of a local or a piecewise (non-local) rule? Since the *CV.CV* pattern is ambiguous in the sense that it can be both SL3 and SP2, we need to test computationally more difficult forms like *CV.CV.CV*. If neural responses to hypothetical words such as *sakasa* vs *sakaʃa* are being tested, we will observe whether participants are responding to a local rule (SL5) or non-local rule (SP2). If the results show behavioral sensitivity, the P3 and LPC to violations of a pattern that has still the same non-local but different local parameters, then it can be concluded that participants are responding to a non-local (SP) rule. If there will be less or more behavioral sensitivity with or without neural responses, then it can be concluded that participants are responding to a local (SL) rule.

Contrary to Experiment 4 which used *CV.CV* words that are both SL3 and SP2, participants of this experiment were tested with words in the form of *CV.CV.CV* that are both SL5 and SP2. By keeping the piecewise feature, the same (SP2), and increasing the computational complexity from SL3 to SL5, I aimed to observe the effect of the degree of locality. The experimental details are explained in section 8.2. The results, presented in section 8.3, showed that participants were not able to learn the pattern behaviorally; and thus, they did not exhibit an ERP response modulated by well-formedness. I discuss the implications of these results in section 8.4 and section 8.5 concludes the chapter by reporting that the degree of locality in the pattern modulates both learning and brain response. I conclude that exposure to computationally different patterns results in both different knowledge states and different underlying neural commitments.

8.1 Methods

8.1.1 Participants

A total of 20 University of Delaware students were recruited as participants and provided written consent for participation in the experiment. Each participant received course credit for participation. 18 of the 20 participants were female. One participant was left-handed. The mean age was 20 (*SD=3.11*, range=18 to 32). None of the participants reported a history of hearing loss or speech/language impairments, and all reported speaking English as their first and only language.

8.1.2 Stimuli

The stimuli used in Experiments 3 and 4 were in the form of *CV.CV*, but the stimuli used in the current experiment were in the form of *CV.CV.CV*. The pattern was the Sibilant Harmony rule again. The stimuli used in this study again were adapted from Lai (2012, 2015). The stimuli were originally recorded from a native speaker of Mandarin Chinese with phonetic training, pronouncing the words in *CV.CV.CVC* format. The third syllable without the coda (CV) was segmented from the sample words using the offset/onset of noise for the surrounding vowel and sibilant as a criterion. I then applied a 10 ms sinusoidal fade-in and fade-out to the beginning and end of each recording to eliminate the effects of trimmed formants. Next, I merged three syllables and constructed the *CV.CV.CV* words. As a final step, the peak amplitude of all items was normalized to their mean. See Figure 8.1 for a spectrogram view from Praat software (Boersma & Weenink, 2017). The duration of each syllable was strictly controlled – the first and third syllables at 200 ms, the second syllable at 250 ms, making each word 650 ms long. The third syllable (marking the violation) began at 450 ms.

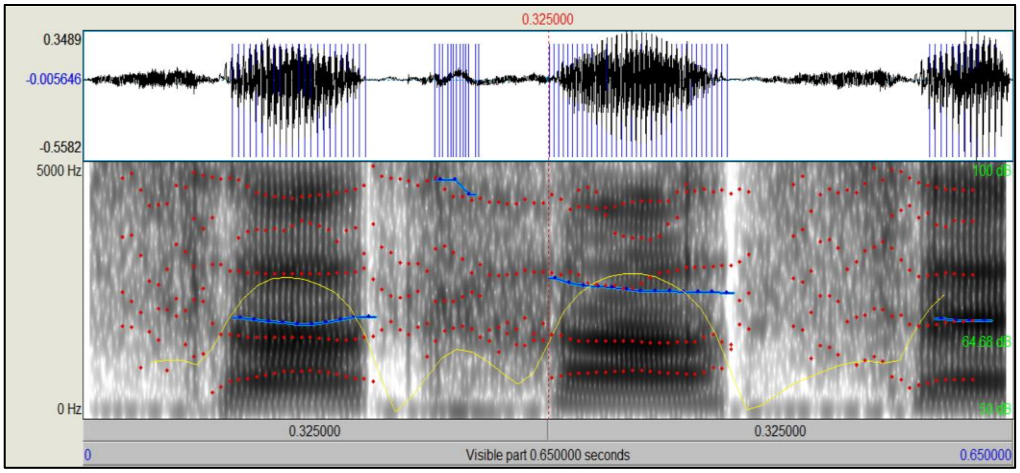


Figure 8.1: Praat spectrogram for an example word, [sakase].

All training and test stimuli consisted of three syllables of the form *CV.CV.CV*, with sibilants ([s, ʃ]) as the first and third consonants, and ([k]) as the second consonant. The vowels in the alphabet of the language were [a, ɛ, ɔ, i, u]. 500 words were constructed with half agreeing in sibilant anteriority (e.g. [sakuso], [ʃekaʃi]), and the other half disagreeing (e.g. [sakeʃi], [ʃekiso]). Of the 250 words that had an agreement, 50 appeared only in the training session. The other 200 were used in the test phase as novel-fit items (together with 50 trained words). All of the 250 words that had a disagreement were used in the test as novel-violate items.

In order to exclude potential artifacts due to phonetic differences, I checked the mean fundamental frequency (Hz) and mean intensity (dB). See Table 8.1 below for the statistical comparisons. It can be seen that with respect to phonetic parameters, there is no difference between grammatical vs. ungrammatical words.

Table 8.1: Phonetic properties of grammatical vs ungrammatical words (mean fundamental frequency and mean intensity) and comparisons between types of words.

Word Type	Sibilant Tier	F_0 (mean)	Intensity (mean)
Grammatical	[s.k.s] or [ʃ.k.ʃ]	263.4 (*SD=5.92*) Hz	72.1 (*SD=0.52*) dB
Ungrammatical	[s.k.ʃ] or [ʃ.k.s]	263.4 (*SD=5.88*) Hz	72.0 (*SD=0.51*) dB
Grammatical vs Ungrammatical		t(249)=.011; p=.991	t(249)=.796; p=.427

8.1.3 Apparatus and procedure

The apparatus was the same as the previous experiments. The procedure for the experiment consisted of two phases: a training phase and a testing phase.

During the training phase, participants listened to 200 tokens (50 grammatical words repeated four times) and were instructed to repeat each word orally after they heard it. The training lasted approximately 15 minutes. In the testing phase, the task for the participant was to press a button in response to each stimulus to categorize the stimulus as grammatical or ungrammatical, according to the rule they had learned in the training phase. Participants responded with the left hand for grammatical words and with the right hand for ungrammatical words. Participants were tested in an auditory oddball paradigm and received a deviant (ungrammatical) stimulus appearing infrequently among repeated occurrences of a standard (grammatical) stimulus.

Participants discriminated between grammatical and ungrammatical words, with grammatical words appearing in 83% of trials (1000 words) and ungrammatical words appearing in 17% of trials (200 words), with 1200 total trials. Each participant was presented with a total of 200 rare stimuli and 1000 frequent stimuli. The stimuli were delivered continuously, with a random number (between 3 and 7) of standards between each deviant. Each word was presented for 650 ms followed by a blank inter-trial interval of 1100 to 1500 ms. The duration of the test session was approximately 40 minutes. No explicit feedback was given to the participants during the test phase. All participants easily understood the instructions.

8.1.4 Data recording and analysis

Behavioral and electrophysiological data recording and analysis were exactly the same as in Experiment 4.

8.2 Results

8.2.1 Behavioral Results

Ungrammatical words were detected with a mean sensitivity of -0.081 (d') (*SD=0.450*), a score not significantly different from zero, t(19)=0.812, p=.427, d=0.181. Mean bias rate was -0.438 (c) (*SD=.30*) – negative and significantly different than zero which is expected and the artifact of the oddball paradigm. See Figure 8.2 below for a visual comparison.

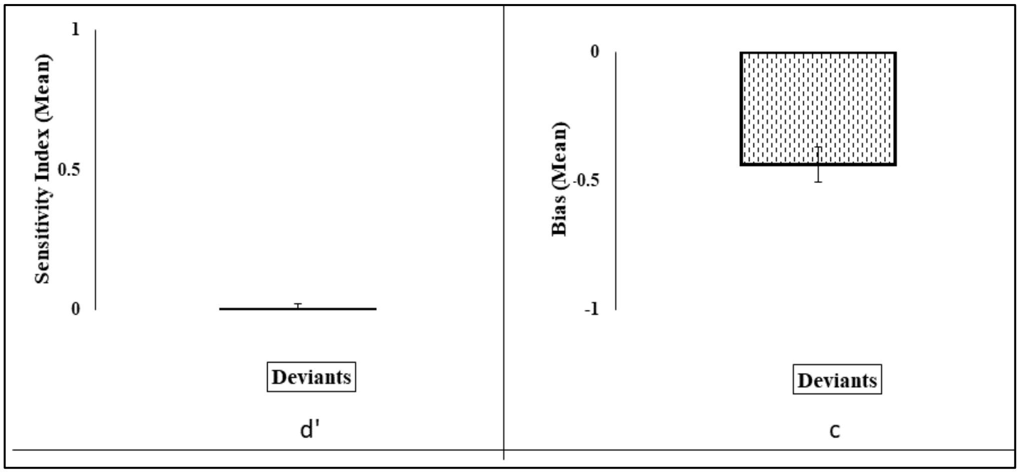


Figure 8.2: Averages of sensitivity index rates (left panel) and bias rates (right panel). Error bars indicate standard error of the mean.

These results indicate that deviants were not detected with a mean sensitivity higher or different than zero which suggests that participants did not learn the rule, even though they were biased towards thinking that most words are grammatical. Mean accuracy (percent correct) for grammatical words was .64 (SD=.15) and .32 (SD=0.10) for ungrammatical. The difference was significant t(19)=6.68, p<.001, d=1.49, 1-β=.999. Mean RT was 474 ms

(SD=122) for grammatical words and 476 ms (SD=126) for ungrammatical. The difference was not significant, t(19)=0.539, p=.596.

8.2.2 P3 Rare-minus-frequent difference waveforms

8.2.2.1 Stimulus-locked waveforms

Figure 8.3 shows stimulus-locked ERP absolute waveforms and the difference waves (rare-minus-frequent). Both frequent and rare stimuli revealed a typical auditory evoked potential (AEP) at frontal-central electrodes. There was a small difference between grammatical (blue) and ungrammatical (red) words beginning from the offset of the word (650ms) to 1000 ms, specifically at the frontal and central regions. However, this difference was not statistically significant.

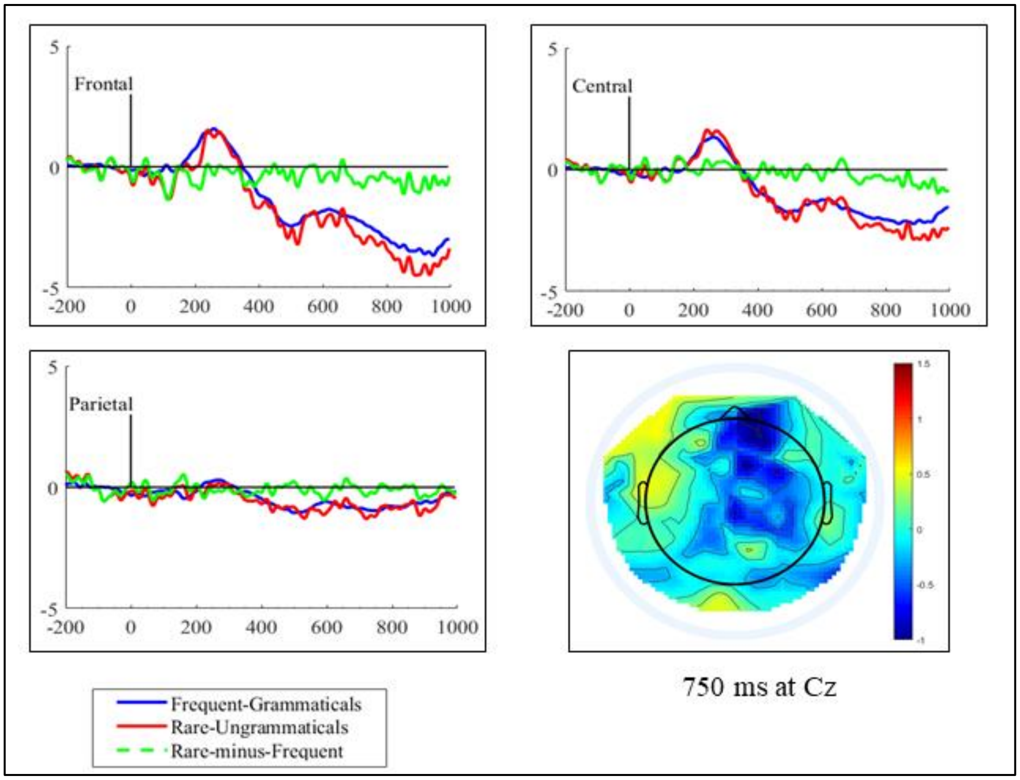


Figure 8.3: Stimulus-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Topoplot shows the mean difference wave at 750 ms at Cz.

ANOVAs revealed a significant region effect but not a grammaticality effect, and there was no interaction. Statistical results are summarized in Table 8.2 below.

Table 8.2: Stim-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

Voltage (μV)			Statistics		
Frontal	Central	Parietal	Gramm. df= 1,19	Region df= 2,38	Region × Gramm. df= 2,40
.267 (.366)	.050 (.281)	.052 (.362)	F=0.161 p=.693 η^2=.008	F=9.429 p<.001 η^2=.332 1-β=0.999	F=0.497 p=.612

8.2.2.2 Response-locked waveforms

Figure 8.4 below shows response-locked ERP absolute waveforms and the difference waves (rare-minus-frequent). Again, a typical Bereitschafts Potential (BP) was observed in the parietal region. Next, at the frontal region, the difference wave was different from zero which can be taken as an indication of differential processing of grammatical vs ungrammatical words. On the contrary, at central and parietal regions, there was no difference before the motor response, and the difference wave bounces around zero.

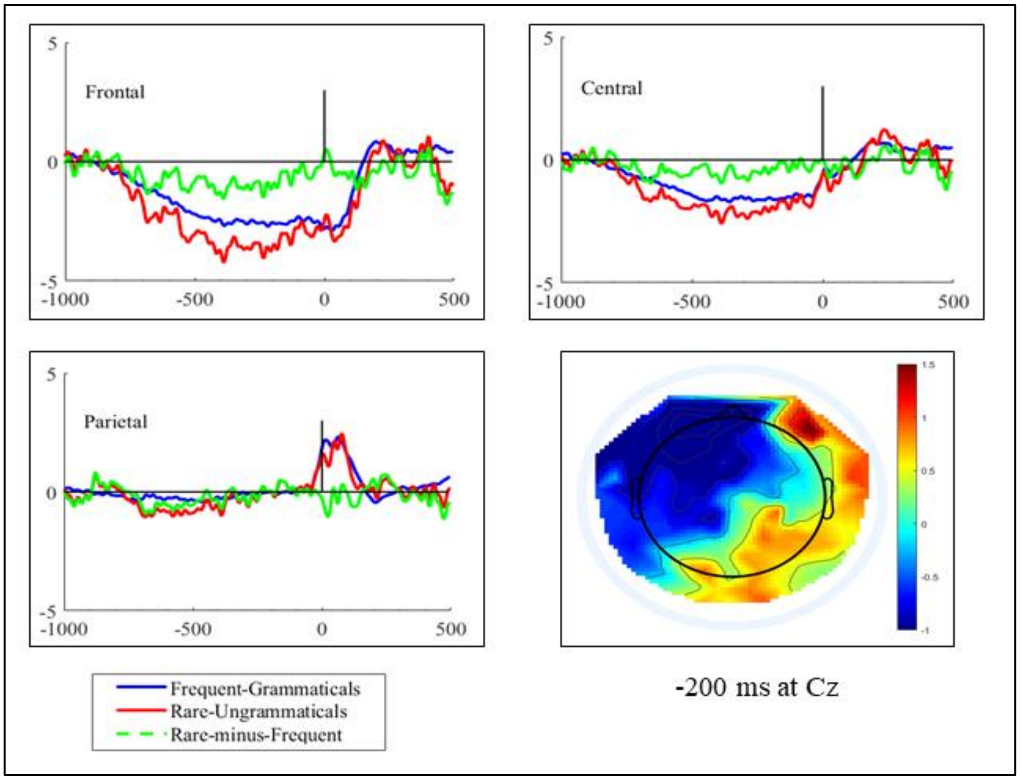


Figure 8.4: Response-locked grand average ERP absolute and difference (rare minus frequent) waveforms at the frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4) midline electrode sites. These waveforms isolate the brain's differential processing of grammatical and ungrammatical stimulus categories. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Topoplot shows the mean difference wave at -200 ms at Cz.

Statistical analyses mirror the stimulus-locked results in that there was a significant region effect, but no grammaticality effect or interaction was observed. See Table 8.3 below for response-locked statistical results.

Table 8.3: Resp-locked rare-minus-frequent difference wave measures (within-subjects standard errors in parentheses), along with F, p, η2, 1-β values for statistical analyses.

Voltage (μV)			Statistics		
Frontal	Central	Parietal	Gramm. df= 1,19	Region df= 2,38	Region × Gramm. df= 2,40
.672	.252	-.191	F=0.598	F=37.441	F=2.038
(.496)	(.337)	(.350)	p=.449	p<.001	p=.166
			η^2=.031	η^2=.663	
				1-β=1.000	

8.2.3 LPC Analysis

Figure 8.5 below shows that neither the grammatical nor the ungrammatical waveforms were positive-going in the very late time window. Topo-plot indicates that the contrast between grammatical and ungrammatical words did not reveal significant parietal positivity effect for non-words. Therefore, no statistically significant difference was observed between the grammatical and ungrammatical words, t(19)=0.126, p=.901, d=.028. The absence of an LPC implies that participants did not detect any anomaly in response to the violation of sibilant harmony rule.

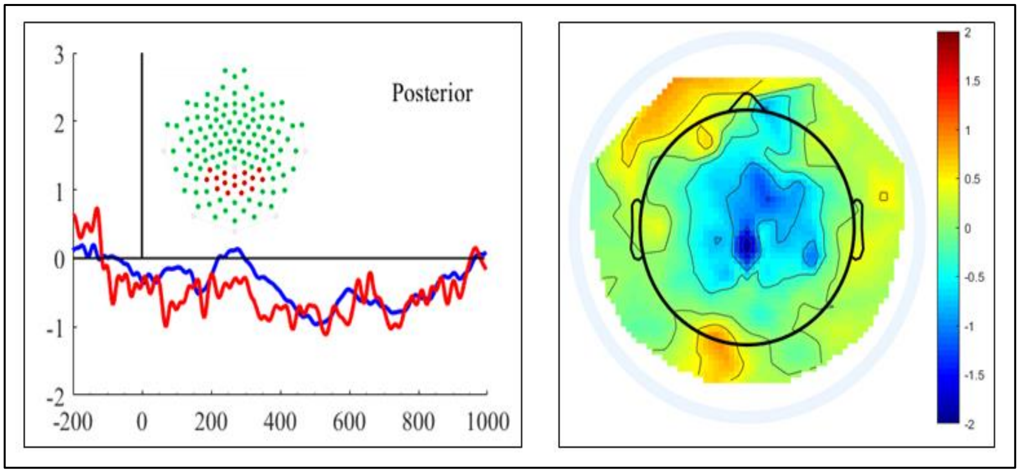


Figure 8.5: Left Panel: Grand average ERP absolute waveforms obtained for grammatical words (blue line) and ungrammatical words (red line) at the posterior region. The x-axis depicts the time (in ms) from the onset of the word, while y-axis depicts voltage (in µV). Right panel: Spatial topography maps showing voltage differences between grammatical and ungrammatical words at 1000 ms.

8.3 Discussion

The aim of this experiment was to look for the neurophysiological markers of a more complex rule-processing mechanism. When the rule or pattern to be learned was an SL3 and SP2 in regard to the subregular hierarchy, the violations of this rule elicited moderate behavioral sensitivity plus robust P3 and LPC at the neural level (Experiment 4). In this chapter, I tried to replicate these results with a computationally more complex pattern. I tested whether the neural responses (P3 and LPC) were indexing violations of a local or a piecewise (non-local) rule. The pattern that was tested in Experiment 4 was SL3 from the point of local formulation and SP2 in terms of non-local formulation. In this current experiment, I tested a new pattern by keeping the non-locality aspect the same; a pattern that is SL5 in terms of the locality. Thus, the hypothesis was that if the behavioral or neural results change, then it can be concluded that it was because of a change at the local level not at the piecewise level. By doing this, I tested whether participants were responding to a local dependency violation or a long-distance dependency violation. Therefore, this experiment was an attempt to replicate Experiment 4 with a computationally different pattern.

The results showed that when the locality aspect changed (from SL3 to SL5), participants did not learn the rule at the behavioral level. Thus, no significant neural response was observed. This result implies that participants were actually responding to a local dependency violation rather than a long-distance dependency in Experiment 4. Since when a pattern that is more complex in terms of locality and the same at the piecewise level, participants showed no behavioral sensitivity. The sensitivity index (*d'*) rate was around zero, which shows no learning at all. Although the accuracy results showed significant differences between the categorization of grammatical vs ungrammatical words, the sensitivity index was not positive. This indicates that participants most of the time missed the signal or gave false alarms. One plausible explanation for the lack of behavioral sensitivity is that when the pattern is more complex (compared to the pattern that was tested in Experiment 4) the categorization task becomes more difficult. This also suggests that the 15 min training was not enough to learn this computationally complex rule.

The absence of behavioral learnability in this experiment where the pattern is both local (SL5) and non-local (SP2) raises another question – specifically, how did participants in Experiment 1 learn the sibilant harmony pattern that is more complex than the harmony pattern in this experiment? The pattern that was tested in Experiment 1 was in the form of CV.CV.CVC; thus, it was both SL7 and SP2. The task for the participants in Experiment 1 was a detection task in which they were asked to press the button only for the words violating the pattern. However, the task for the participants in the current experiment was a categorization task in which they were asked to categorize each word by pressing two different buttons. Therefore, the task was more demanding in this current experiment which may have affected the learnability.

Neurophysiological results also followed the behavioral results. There was no difference between the grammatical vs ungrammatical words in the P3 difference waveform or the LPC. Although there was a significant difference between the different regions, there was not any significant grammaticality effect. These results together with behavioral results showed that when participants were not able to extract the pattern during the training session, they showed no behavioral or neural sensitivity to pattern violations.

8.4 Conclusion

The aim of this chapter was to examine the neurophysiological correlates of learning of a more complex phonotactic pattern. While recording EEG, participants were presented with words that were either well-formed or ill-formed according to the rule. It was found that learners did not show sensitivity indicating knowledge of the rule. In addition, no significant ERP response was found which could have been interpreted as reflecting categorization decision and anomaly detection. These results, together with the results from Chapter 5, show that the neurophysiological response reflecting the functioning of the phonological processing mechanism depends on the computational complexity of the pattern to be learned. Thus, different local patterns converge on different knowledge states with different underlying neural mechanisms.

Chapter 9

GENERAL DISCUSSION AND CONCLUSION

This dissertation explored the neurophysiological correlates of phonotactic pattern learning. It investigated the research question of whether the generalizations acquired in laboratory settings are cognitively encoded. A series of artificial language learning experiments were conducted while recording EEG to test the learnability of non-adjacent phonotactic dependencies and the types of locality relations that are observed in such patterns. The primary goal was to provide new knowledge about whether linguistic constraints that are learned in laboratory situations are directly "channeled" into incremental, real-time phonological predictive processing, and whether this process depends on the type of learning mechanism (domain-specific vs. domain-general) and the computational complexity of the patterns. In particular, this goal was achieved by examining the nature of the neural commitment produced by language exposure.

In Experiment 1, the learnability assumptions of the Subregular Complexity Hypothesis (Heinz, 2010) were tested to see whether the learnability of specific phonotactic patterns is restricted to particular subregular regions of the complexity hierarchy. It has been shown that participants showed significantly greater sensitivity to a pattern that is constrained by the Subregular Hierarchy. Therefore, the pattern that resides inside of specific subregular regions was easily learned. However, sensitivity levels for the unattested and predicted-to-be-unlearnable patterns also showed evidence of learning. One plausible explanation is that innate linguistic factors are operative during learning, added on top of general psychological learning/problem-solving mechanisms. In this way, the Subregular Hypothesis can be thought of an example of a domain-specific constraint on induction such that patterns which are attested in human languages exhibit constrained learning because they are channeled to language-specific learning modules. It has been concluded that a dual mechanism of phonological processing is active when parsing phonotactic patterns. There is the linguistic domain-specific mechanism which is constrained by the Subregular Hypothesis in that patterns with specific subregular computational properties are privileged with respect to learnability. And there is the psychological domain-general mechanism which is not constrained by the Subregular Hypothesis and helps the unattested patterns be learned in a laboratory setting.

In Experiments 2 and 3, I conducted two different EEG experiments with different hypotheses about the elicitation of different ERP components (MMN and LPC). The hypothesis behind these experiments was that if knowledge of phonological patterns is part of an internal grammar governing language perception and production, then these features of the grammar should extend into the perceptual system which should be observable with ERPs. Although behavioral results showed evidence of learnability, EEG results showed no significant neurophysiological marker of this learning. It was discussed that since the MMN is considered to be a limited perceptual response and does not have access to higher-order regularities, it was not able to reflect the internal phonotactic grammar features. This implies that the early auditory processing stage may not be sensitive to non-adjacent patterns. However, in the absence of significant neural responses, this is just a prediction for future studies. In addition, it was discussed that the computational complexity of the pattern that was tested might be the reason for not observing an LPC.

In Experiment 4, after discussing that the MMN does not index phonotactic pattern violations and the LPC was not elicited due to the computational complexity of the pattern, the phonotactic pattern was simplified. After the simplification of the pattern, the aim of the next experiment was focused more on finding the neurophysiological marker of implicit phonological rule learning rather than studying the neurophysiological markers of the attested vs unattested patterns. The hypothesis was to incorporate rule knowledge in a categorization task and look for P3 difference waveform as an index of categorization differences between harmonic and disharmonic words, in addition to the LPC as an index of anomaly detection in a rule-governed phonological sequence. The results showed that the peak of the P3 difference wave appeared at 300 ms after the violation point. This peak indicated the point in time when the participant categorized the stimulus as grammatical or ungrammatical, prior to the behavioral response (one needs to determine grammaticality first to choose a response selection). This implies that the brain detected that a particular stimulus was violating the pattern exactly at the violation point, not after the whole word had been processed. In other words, the phonological pattern detection mechanism does not wait until the end of the word to rule out an ungrammatical word. This time course information cannot be inferred from behavioral measures alone. These results were also backed up by the response-locked waveforms that showed a positive peak immediately preceding the response. This wave reflects the decision activity between the categorization and the button press response, which consists of selecting, preparing, and executing the appropriate response.

In addition to the P3 reflecting a categorization difference, LPC modulations indicated that phonotactic constraint violations affected later stages of cognitive processing. This was important because the LPC was elicited in response to phonotactic violations even when ungrammatical words had already been categorized as violating the pattern. Therefore, LPC elicitation adds something more valuable on top of P3 which is that the LPC is reflecting anomaly detection in rule-governed phonological sequences. This can be taken as an indication of the instantiation of rule-based knowledge into the learner’s real-time language processing system.

In Experiment 5, after it had been shown that phonotactic violations were reflected by the P3 difference waveform as a marker of categorization and LPC as indexing phonological anomaly detection, I decided to replicate these results with a different learning mechanism, using explicit instruction. The aim of this experiment was to look for the neurophysiological marker of the explicit phonological rule-learning mechanism. As investigating the differences between implicit and explicit phonotactic learning is important to see whether different learning processes are encoded differently at the neural level. Behavioral results showed that participants’ sensitivity to the ungrammatical words was very high – specifically, the explicit group was much more sensitive to the violation of phonotactic rule than the implicit group in Experiment 3. On the contrary, the neural results showed the reverse. The explicit learners showed no P3 or LPC modulation, despite the presence of a robust AEP in stimulus-locked waveforms and Bereitschafts Potential in response-locked waveforms. Both stimulus-locked and response-locked P3 difference waveforms showed that grammatical and ungrammatical words were processed similarly yielding no significant grammaticality effect. These results were interpreted to indicate that implicit and explicit learning leads to different types of neural encoding of the acquired phonotactic rule. It can be concluded that implicit learning leads to a measurable neural learning response typical of the categorization and anomaly detection systems, whereas explicit learning leaves the brain silent.

Finally, Experiment 6 was conducted to test the computational complexity effect on the neural system, specifically to see whether the P3 and LPC are indexing violations of a local or a piecewise (non-local) pattern. The results showed that learners did not show behavioral sensitivity indicating knowledge of the rule. In addition, no significant ERP response was found which could have been interpreted as reflecting categorization decision or anomaly detection, even though the AEP and BP were elicited. These results showed that the neurophysiological response reflecting the functioning of the phonological processing mechanism depends on the computational complexity of the pattern to be learned. Thus, different local patterns converge on different knowledge states with different underlying neural mechanisms.

The primary goal of this dissertation was to provide new knowledge about whether linguistic constraints that are learned in laboratory situations are directly “channeled” into incremental, real-time phonological predictive processing, and whether this process depends on the type of learning mechanisms (domain-specific vs. domain-general) and the computational complexity of the patterns. This dissertation has made the following contributions: Firstly, human phonological learning mechanisms are restricted mainly by domain-specific computational constraints that are supported by domain-general mechanisms. Second, phonological constraints that are acquired in laboratory situations are directly channeled into incremental, real-time phonological predictive processing evidenced by millisecond level sensitive ERP components. Third, implicit learning, a variety of domain-specific learning mechanisms, and explicit learning, a variety of domain-general learning mechanisms, converge on similar knowledge states, but with different underlying neural mechanisms. And finally, the nature of the neural commitment and the underlying knowledge state depends on the type of learning and the computational complexity of the patterns.

9.1 Future Studies

When I started working on this research topic, my aim was to present some answers to the questions that I raised in the introduction chapter of this dissertation. After five years of research and more than two hundred pages of a dissertation, I think I caused more questions than simple answers. Nevertheless, future work that addresses new questions raised in the above should be continued. In this section, I will enumerate briefly which experiments could be done in the future based on what I found so far.

The number one thing that future psycholinguists or cognitive neuroscientists should keep in mind is to start with a simple experiment. An experiment that partially replicates a previous finding, and builds on the previous findings. In this respect, based on the results of Aaltonen et al. (2008), the first experiment should look for MMN in response to sibilant harmony violation with the short *C.V.C.V* words without any training. A language whose alphabet has only one vowel /a/ and two consonants /s, ʃ/ can be used. The standard stimuli in the oddball design would be [sasa] and [ʃaʃa] and the deviants would be [saʃa] and [ʃasa]. My prediction is that this experiment will definitely yield a MMN. However, I am not sure what would this MMN response tell us? It can be due to rule violation or just a simple sound change. A follow-up experiment can answer this question by including an intervening CV like in [sakasa]. Depending on which ERP is elicited (P3 or MMN), it can be discussed whether MMN is indexing pattern violation or just a local sound change. The reason I kept the vowel just the same is that most of my participants told me informally after the experiment that whatever rule they had in their mind was about vowels. That is, many participants during the training were trying to extract a rule depending on the vowels which of course was impossible because vowels were presented randomly.

Another simple experiment idea is to test to the effect of sound change in the coda position compared to onset position. The experiment stimuli can have [as] or [aʃ] as standard stimuli and [sa] or [ʃa] as deviant and compare it [sa] or [ʃa] as standard stimuli and [as] or [aʃ] as deviant. A possible MMN amplitude or latency difference would imply the position effect of coda vs onset. This will inform phonological theory about the brain representation of position effects. A simple follow-up experiment would then test [sas] vs [saʃ] and compare the MMN to [ʃaʃ] vs [ʃas].

If the preference is not choosing a simple design, then the next step would be to test the SL5 SP2 pattern (the pattern I tested in Experiment 6) within explicit learning settings. This will complete at least the 2 x 2 statistical design that I initiated: group (implicit vs explicit) and computational complexity of pattern (SL3 vs SL5). With this design, some questions about the effects of explicit instruction on pattern learning can be answered. These sets of experiments can be run by varying the vowels as I did in this dissertation or just using one vowel as discussed in the previous paragraph. A possible extension would also include testing subjects one week later to see whether implicit or explicit learning is more stable. My prediction is that one week later implicit learners will have better sensitivity and better neural commitment. One can also examine how sleep-related consolidation processes contribute to different aspects of pattern learning by just testing the learning after some sleep, possibly one day.

Of course, using neuroimaging techniques to see whether explicit and implicit learning mechanisms are reflected in the brain as distinct neural structures with different connectivity would also be the next step. This way finer differences between the domain-specific vs domain-general learning mechanisms could be discovered, and that may be where the answers to these questions reside.

An additional set of experiments, which was also conducted by Lai (2012) in her dissertation, would be examining whether non-linguistic stimulus patterns that can be described by the same grammars are linguistically constrained even when the stimuli do not bear any resemblance to linguistic forms. The domain-specific hypothesis is that non-linguistic stimuli get channeled to non-linguistic learning mechanisms. On the other hand, patterns that are identified as “linguistic” by the auditory processing system exhibit constrained learning, because they are “channeled” to language-specific learning modules. The stimuli will be non-linguistic melodic tone patterns that follow each grammatical pattern rather than contiguous phonemes as part of words. It is important to stress here that the stimuli that conform to the same rules will be tested from a formal perspective only. I.e., the tone sequences are obviously not words with contiguous phonemes. Lai (2012) converted vowels to cymbals, sibilants to tom drum, and the stop [k] to a snare drum, the duration of each sound was 200 ms and 1400 for seven sounds. Piano tones can also be used for each of the three categories above and the durations will be exactly the same with linguistic stimuli. Using tone sequences will allow us to strictly control the temporal dynamics of the stimuli, and avoid variability in the stimuli that would add noise to the EEG, while at the same time keeping the general patterns/rules the same. The results should generalize to the same pattern of contiguous phonemes in a word.

In conclusion, future work in the context of cognitive neuroscience of phonology is desperately needed.

REFERENCES

Aaltonen, O., Hellström, A., Peltola, M. S., Savela, J., Tamminen, H., & Lehtola, H. (2008). Brain responses reveal hardwired detection of native-language rule violations. *Neuroscience Letters*, *444*(1), 56–59.

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing*, *16*(1), 38–51.

Amenedo, E., & Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *European Journal of Neuroscience*, *12*(7), 2570–2574.

Applegate, R. B. (1972). *Ineseno Chumash grammar*. Ph.D. dissertation, University of California, Berkeley.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annu. Rev. Psychol.*, *56*, 149–178.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, *9*(4), 321–324.

Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition*, *132*(2), 137–150.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677.

Bendixen, A., Prinz, W., Horváth, J., Trujillo-Barreto, N. J., & Schröger, E. (2008). Rapid extraction of auditory feature contingencies. *NeuroImage*, *41*(3), 1111–1119.

Bendixen, A., Schröger, E., Ritter, W., & Winkler, I. (2012). Regularity Extraction from Non-Adjacent Sounds. *Frontiers in Psychology*, *3*(May), 1–12.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, *104*(3), 591–630.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants’ expectations for count nouns and adjectives. *Journal of Cognition and Development*, *4*(3), 357–381.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*(March), 1–15.

Campbell, K. L., & Tyler, L. K. (2018). Language-related domain-specific and domain-general systems in the human brain. *Current Opinion in Behavioral Sciences*, *21*, 132–137.

Carson, J. (1988). Unification and Transduction in Computational Phonology. In *International Conference on Computational Linguistics* (pp. 106–111).

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69--B77.

Chandlee, J. (2014). *Strictly Local Phonological Processes*. Ph.D. dissertation, University of Delaware, Newark.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibanez, A., … Bekinschtein, T. A. (2013). Expectation and Attention in Hierarchical Auditory Prediction. *Journal of Neuroscience*, *33*(27), 11194–11205.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*(5), 351.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*, 113–124.

Cleeremans, A., Schreiber, S. D., & Mcclelland, J. L. (1989). Finite-State Automata and Simple Recurrent Networks. *Neural Computation*, *1*, 372–381.

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The Importance of Sound for Cognitive Sequencing Abilities. *Current Directions in Psychological Science*, *18*(5), 275–279.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21–58.

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*(2), 131–143.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: a study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1355.

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*(1), 138–145.

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*(5), 665–678.

Domahs, U., Kehrein, W., Knaus, J., Wiese, R., & Schlesewsky, M. (2009). Event-related potentials reflecting the processing of phonological constraint violations. *Language and Speech*, *52*(4), 415–435.

Donchin, E. (1981). Surprise!… surprise? *Psychophysiology*, *18*(5), 493–513.

Dröge, A., Fleischer, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2016). Neural mechanisms of sentence comprehension based on predictive processes and decision certainty: electrophysiological evidence from non-canonical linearizations in a flexible word order language. *Brain Research*, *1633*, 149–166.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467.

Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology*, *14*(1–2), 1–52.

Endress, A. D., & Mehler, J. (2010). Perceptual constraints in phonotactic learning. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(1), 235–250.

Eschman, A., St. James, J., Schneider, W., & Zuccolotto, A. (2005). PsychMate: Providing psychology majors the tools to do real experiments and learn empirical methods. *Behavior Research Methods*, *37*(2), 301–311.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. C. M. (2016). The relationship between artificial and second language learning. *Cognitive Science*, *40*(4), 822–847.

Fabiani, M., Gratton, G., & Coles, M. (2000). Event-related brain potentials: methods, theory. *Handbook of Psychophysiology*, 53–84.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84.

Finley, S. (2011). The privileged status of locality in consonant harmony. *Journal of Memory and Language*, *65*(1), 74–83.

Finley, S. (2012). Testing the limits of long-distance learning: Learning beyond a three-segment window. *Cognitive Science*, *36*(4), 740–756.

Finley, S. (2015). Learning nonadjacent dependencies in phonology: Transparent vowels in vowel harmony. *Language*, *91*(1), 48–72.

Finley, S. (2017). Locality and harmony: Perspectives from artificial grammar learning. *Linguistics and Language Compass*, *11*(1), 1–16.

Finley, S., & Badecker, W. (2008). Substantive biases for vowel harmony languages. In *In Proceedings of the West Coast Conference on Formal Linguistics 27*.

Finley, S., & Badecker, W. (2009a). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437.

Finley, S., & Badecker, W. (2009b). Right-to-left biases for vowel harmony: Evidence from artificial grammar. In *Proceedings of the 38th North East linguistic society annual meeting* (Vol. 1, pp. 269–282).

Finn, A. S., & Kam, C. L. H. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*(2), 477–499.

Fitzgerald, P. G., & Picton, T. W. (1984). The effects of probability and discriminability on the evoked potentials to unpredictable stimulia. *Annals of the New York Academy of Sciences*, *425*(1), 199–203.

Folia, V., Uddén, J., de Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial Language Learning in Adults and Children Corresponding author. *Language Learning*, *60*(December), 188–220.

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, *9*(1), 1–24.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*(3), 259–281.

Friederici, A. D., & Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain Research*, *1000*(1–2), 72–77.

Friedman, D., & Johnson, R. J. (2000). Event-Related Poteltnai (ERP) Studies of Memory Encoding and Retrieval: A Selective Review, *28*(January), 6–28.

Frisch, S., Kotz, S. A., von Cramon, D. Y., & Friederici, A. D. (2003). Why the P600 is not just a P300: The role of the basal ganglia. *Clinical Neurophysiology*, *114*(2), 336–340.

Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., & Griffiths, T. D. (2011). Neural prediction of higher-order auditory sequence statistics. *NeuroImage*, *54*(3), 2267–2277.

Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language*, *51*(4), 586–603.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, *7*(5), 567–580.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*(6), 660–676.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive*

Processes, *8*(4), 439–483.

Handy, T. C. (2005). *Event-related potentials: A methods handbook*. MIT press.

Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech*, *1*(1), 1–7.

Harrison, M. A. (1978). *Introduction to Formal Language Theory* (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Heinz, J. (2010). Learning Long-Distance Phonotactics. *Linguistic Inquiry*, *41*(4), 623–661.

Heinz, J. (2018). The computational nature of phonological generalizations. In L. Hyman & F. Plank (Eds.), *Phonological Typology, Phonetics and Phonology* (pp. 126–195). Mouton.

Heinz, J., & Rogers, J. (2013). Learning Subregular Classes of Languages with Factored Deterministic Automata. In A. Kornai & M. Kuhlmann (Eds.), *Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13)* (pp. 64–71). Sofia, Bulgaria: Association for Computational Linguistics.

Hilgard, J., Weinberg, A., Hajcak Proudfit, G., & Bartholow, B. D. (2014). The Negativity Bias in Affective Picture Processing Depends on Top-Down and Bottom-Up Motivational Significance. *Emotion*, *14*(5), 940–949.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281–312.

Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2006). *Introduction to Automata Theory, Languages, and Computation (3rd Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Huotilainen, M., Ilmoniemi, R. J., Lavikainen, J., Tiitinen, H., Alho, K., Sinkkonen, J., … Näätänen, R. (1993). Interaction between representations of different features of auditory sensory memory. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*.

Hwangbo, H. J. (2015). Learnability of two vowel harmony patterns with neutral vowels. In *Proceedings of the The third Annual Meeting on Phonology (AMP 2015)*.

Jahanshahi, M., & Hallett, M. (2003). *The Bereitschaftspotential: Movement-related cortical potentials*. Springer Science & Business Media.

Johnson, C. D. (1972). *Formal Aspects of Phonological Description*. The Hague: Mouton.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, *30*(1), 90–97.

Johnson, R. (1995). Event-related potential insights into the neurobiology of memory systems. *Handbook of Neuropsychology*, *10*, 135.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66.

Kane, N. M., Curry, S. H., Butler, S. R., & Cummins, B. H. (1993). Electrophysiological indicator of awakening from coma. *The Lancet*, *341*(8846), 688.

Kaplan, R. M., & Kay, M. (1994). Regular Models of Phonological Rule Systems. *Comput. Linguist.*, *20*(3), 331–378.

Keil, A., Debener, S., Gratton, G., Junghofer, M., Kappenman, E. S., Luck, S. J., … Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*(1), 1–21.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*(2), 201–258.

Kobele, G. M. (2006). *Generating Copies : An investigation into structural identity in language and grammar*. Ph.D. dissertation, University of California, Los Angeles.

Koo, H., & Callahan, L. (2012). Tier-adjacency is not a necessary condition for learning phonotactic dependencies. *Language and Cognitive Processes*, *27*(10), 1425–1432.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1985). Processing of stimulus properties: evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 393.

Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science, 273*(5277), 971–973.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 979–1000.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, *197*(4305), 792–795.

Lai, R. (2012). *Domain specificity in learning phonology*. Ph.D. dissertation, University of Delaware, Newark.

Lai, R. (2015). Learnable vs. Unlearnable Harmony Patterns. *Linguistic Inquiry*, *46*(3), 425–451.

Lei, Y., Wang, J., Zhu, Y., Chen, Q., & Li, H. (2019). P300 and positive slow waves reveal the plausibility in inductive reasoning. *Psychophysiology*, e13337.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, *12*(4), 433–452.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75.

Linzen, T., & Gallagher, G. (2014). The timecourse of generalization in phonotactic learning. In *Proceedings of the Annual Meetings on Phonology* (Vol. 1).

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Kappenman, E. S. (2011). *The Oxford handbook of event-related potential components*. Oxford university press.

Luck, S. J., Kappenman, E. S., Fuller, R. L., Robinson, B., Summerfelt, A., & Gold, J. M. (2009). Impaired response selection in schizophrenia: evidence from the P3 wave and the lateralized readiness potential. *Psychophysiology*, *46*(4), 776–786.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, *66*(3), 309–332.

Magliero, A., Bashore, T. R., Coles, M. G. H., & Donchin, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology*, *21*(2), 171–186.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77–80.

Martin, A., & Peperkamp, S. (2011). Speech perception and phonology. *The Blackwell Companion to Phonology*, *4*, 2334–2356.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379–413.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101--B111.

McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruites on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *59*(6), 477–488.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, *60*(SUPPL. 2), 123–150.

McNaughton, R., & Papert, S. (1971). *Counter-Free Automata*. MIT Press.

Menning, H., Roberts, L. E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport*, *11*(4), 817–822.

Moore-cantwell, C., Pater, J., Staubs, R., Zobel, B., & Sanders, L. (submitted). Event-related potential evidence of abstract phonological learning in the laboratory. Available online at: http://www.clairemoorecantwell.org/phonolearn/moore-cantwell-et-al-2018-ERP-ALL.pdf

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, *25*(1), 83–127.

Moreton, E. (2012). Inter-and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language*, *67*(1), 165–183.

Moreton, E., & Pater, J. (2012a). Structure and substance in artificial-phonology learning, part I: Structure. *Language and Linguistics Compass*, *6*(11), 686–701.

Moreton, E., & Pater, J. (2012b). Structure and substance in artificial-phonology learning, part II: Substance. *Language and Linguistics Compass*, *6*(11), 702–718.

Moreton, E., Pater, J., & Pertsova, K. (2017). Phonological Concept Learning. *Cognitive Science*, *41*(1), 4–69.

Moreton, E., & Pertsova, K. (2016). Implicit and Explicit Processes in Motor Learning. In J. S. and D. Waughtal (Ed.), *Proceedings of the 40th annual Boston University Conference on Language Development* (pp. 277–290). Somerville, MA: Cascadilla Press.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*(4), 933–947.

Musiek, F. E., & Baran, J. A. (2007). The Auditory System: Anatomy. *Physiology and Clinical Correlates*.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller, C. (2003). Broca’s area and the language instinct. *Nature Neuroscience*, *6*(7), 774–781.

Näätänen, R. (1985). Selective attention and stimulus processing: Reflections in event-related potentials, magnetoencephalogram and regional cerebral blood flow. *Attention and Performance*, *11*, 355–373.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*(1), 1–21.

Näätänen, Risto, Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., … others. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432.

Näätänen, Risto, & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. *Biological Psychology*, *8*(2), 81–136.

Näätänen, Risto, Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, *98*(2), 217–221.

Näätänen, Risto, Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590.

Näätänen, Risto, Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence'in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283–288.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162.

Núñez-Peña, M. I., & Honrubia-Serrano, M. L. (2004). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research*, *18*(2), 130–141.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13--B23.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*(2), 225–237.

Osterhout, L. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas

(1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 14*(1), 1–14.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786.

Öttl, B., Jäger, G., & Kaup, B. (2015). Does formal complexity reflect cognitive complexity? Investigating aspects of the Chomsky hierarchy in an artificial language learning study. *PloS One*, *10*(4), e0123059.

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, *265*(3), 179–182.

Paavilainen, P., Jiang, D., Lavikainen, J., & Näätänen, R. (1993). Stimulus duration and the sensory memory trace: an event-related potential study. *Biological Psychology*, *35*(2), 139–152.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733.

Pater, J. (2014, October). Implicit and explicit learning. Retrieved from https://blogs.umass.edu/comphon/2014/10/09/implicit-and-explicit-learning/

Peperkamp, S., & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. *Laboratory Phonology*, *9*, 315–338.

Peperkamp, S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, *101*(3), B31--B41.

Peperkamp S., S. K. & D. E. (2006). The role of phonetic naturalness in phonological rule acquisition. In *Proceedings of the 40th Boston University Conference on Language Development, Somerville, Mass., pp. TBA. Cascadilla* (Vol. 30, pp. 464–475).

Pinker, S. (2013). *Learnability and cognition: The acquisition of argument structure*. MIT press.

Plante, E., Gomez, R., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, *35*(5), 453–462.

Polich, J. (1990). P300, probability, and interstimulus interval. *Psychophysiology*, *27*(4), 396–403.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Pycha, A., Nowak, P., Shin, E., & Shosted, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. In *Proceedings of the 22nd west coast conference on formal linguistics* (Vol. 22, pp. 423–435).

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863.

Reber, R., & Perruchet, P. (2003). The use of control groups in artificial grammar learning. *The Quarterly Journal of Experimental Psychology Section A*, *56*(1), 97–115.

Rebuschat, P. (2015). *Implicit and explicit learning of languages*. *John Benjamins Publishing Company* (Vol. 48).

Rogers, J., Heinz, J., Bailey, G., Edlefsen, M., Visscher, M., Wellcome, D., & Wibel, S. (2010). On Languages Piecewise Testable in the Strict Sense. In C. Ebert, G. Jäger, & J. Michaelis (Eds.), *The Mathematics of Language* (Vol. 6149, pp. 255–265). Springer.

Rogers, J., & Pullum, G. K. (2011). Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information*, *20*(3), 329–342.

Rugg, M. D. (1995). Memory and consciousness: A selective review of issues and data. *Neuropsychologia*, *33*(9), 1131–1141.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition.* Oxford University Press.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595.

Saarinen, J., Paavilainen, P., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport*.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928.

Sallinen, M., Kaartinen, J., & Lyytinen, H. (1994). Is the appearance of mismatch negativity during stage 2 sleep related to the elicitation of K-complex? *Electroencephalography and Clinical Neurophysiology*, *91*(2), 140–148.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *62*(6), 437–448.

Sapir, E., & Hoijer, H. (1967). *The phonology and morphology of the Navaho language* (Vol. 50). University of California Press.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, *1*(3–4), 289–316.

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the Acoustical Society of America*, *106*(2), 1078–1083.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., … Näätänen, R. (2002). Abstract phoneme representations in the left temporal cortex: magnetic mismatch negativity study. *Neuroreport*, *13*(14), 1813–1816.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1–2), 39–91.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568.

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience*, *5*(1), 1–13.

Soli, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *The Journal of the Acoustical Society of America*, *70*(4), 976–984.

Somers, M., Aukes, M. F., Ophoff, R. A., Boks, M. P., Fleer, W., de Visser, K. L., … Sommer, I. E. (2015). On the relationship between degree of hand-preference and degree of language lateralization. *Brain and Language*, *144*(0), 10–15.

St. James, J. D. (2006). *Exploring research methods in psychology using PsychMate*. Pittsburgh, PA: Psychology Software Tools, Inc.

Sussman, E. S., Chen, S., Sussman-Fort, J., & Dinces, E. (2014). The five myths of MMN: redefining how to use MMN in basic and clinical research. *Brain Topography*, *27*(4), 553–564.

Sussman, E., & Winkler, I. (2001). Dynamic sensory updating in the auditory system. *Cognitive Brain Research, 12*, 431–439.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.

Szaflarski, J. P., Rajagopal, A., Altaye, M., Byars, A. W., Jacola, L., Schmithorst, V. J., … Holland, S. K. (2011). Left-handedness and language lateralization in children. *Brain Research, 1433*, 85–97.

Tanner, D., Grey, S., & van Hell, J. G. (2017). Dissociating retrieval interference and reanalysis in the P600 during sentence comprehension. *Psychophysiology, 54*(2), 248–259.

Tervaniemi, M., Maury, S., & Näätänen, R. (1994). Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity.pdf. *NeuroReport*, 844–846.

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning & Memory*, *8*(5), 295–300.

Tessier, A.-M. (2006). Testing for OO-Faithfulness in artificial phonological acquisition. In *Proceedings of the 40th Boston University Conference on Language Development, Somerville, Mass., pp. TBA. Cascadilla* (pp. 619–630).

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, *372*(6501), 90.

Tremblay, K, Kraus, N., & McGee, T. (1998). The time-course of auditory perceptual learning: Which comes first, the chicken or the egg. *NeuroReport*, *9*, 3557–3560.

Tremblay, Kelly, Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: generalization to novel stimuli following listening training. *The Journal of the Acoustical Society of America*, *102*(6), 3762–3773.

Truckenbrodt, H., Steinberg, J., Jacobsen, T. K., & Jacobsen, T. (2014). Evidence for the role of German final devoicing in pre-attentive speech processing: A mismatch negativity study. *Frontiers in Psychology*, *5*(NOV), 1–11.

Tse, C. Y., Low, K. A., Fabiani, M., & Gratton, G. (2012). Rules rule! brain activity dissociates the representations of stimulus contingencies with varying levels of complexity. *Journal of Cognitive Neuroscience*, *24*(9), 1941–1959.

Van Hell, J. G., Fernandez, C. B., Kootstra, G. J., Litcofsky, K. A., & Ting, C. Y. (2018). Electrophysiological and experimental- behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism*, *8(1)*(June), 134-161.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190.

Wacongne, C., Labyt, E., Wassenhove, V. van, Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754–20759.

Wheeldon, L. R., Meyer, A. S., & Smith, M. (2005). Language Production, Incremental. In L. Nadel (Ed.), *Encyclopedia of cognitive science*.

Hoboken, NJ: Wiley.

White, J., & Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, *133*(1), 85–90.

Wilson, C. (2003). Experimental investigation of phonological naturalness. In *Proceedings of the 22nd west coast conference on formal linguistics* (Vol. 22, pp. 533–546).

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, *30*(5), 945–982.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3–4), 147–163.

Winkler, I., Cowan, N., Csépe, V., Czigler, I., & Näätänen, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *Journal of Cognitive Neuroscience*, *8*(5), 403–415.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*(1–2), 239–252.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., … Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*(5), 638–642.

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., … others. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*(3), 357–369.

Yang, J., & Li, P. (2012). Brain Networks of Explicit and Implicit Learning. *PLoS ONE*, *7*(8).

Zaccarella, E., & Friederici, A. D. (2015). Reflections of word processing in the insular cortex: a sub-regional parcellation based functional assessment. *Brain and Language*, *142*, 1–7.

Appendix A

EXPERIMENT 1: TRAINING STIMULI

Conditions	Stimuli			
Sibilant Harmony	sakesas	sisekos	shakushish	shishokesh
	sakisos	sisokos	shakushush	shishukosh
	sasakus	sokosis	shashekush	shokashesh
	sasokos	sokusas	shashokush	shokeshish
	sekeses	sosakos	shekeshosh	shoshokesh
	sekoses	sosikas	shekeshush	shoshokush
	sesukis	sukasus	sheshokish	shukoshesh
	sesukos	sukesas	sheshukush	shukoshosh
	sikasos	susokis	shikeshash	shushakesh
	sikisis	susukos	shikishosh	shushakish
First-Last Assimilation	sakesas	sikasos	shakasosh	sheshokish
	sakeshes	sikushus	shakushish	shikeshash
	sashakus	sisekos	shashekush	shikisash
	sasokos	sishokos	shasokash	shokisosh
	sekeses	sokoshos	shasokush	shoshokush
	sekeshes	sokosis	shekasish	shukesish
	seshokos	sosakos	shekeshosh	shukoshosh
	seshukos	soshakos	shishokesh	shusakesh
	sesukos	sukasus	shisikash	shusakish
	sikashos	susokis	shokashesh	shushakish
Intensive First-Last Assimilation	sakeshes	seshukis	shakasosh	shikisash
	sakishos	seshukos	shakosash	shisikash
	sakushes	sikashos	shasakesh	shisokesh
	sashakus	sikushas	shasakish	shokesish
	sashekes	sikushus	shasekesh	shokisosh
	sashokos	sishekos	shasekish	shukesish
	sashukes	sishokos	shasokash	shukosesh
	sekeshes	sokashas	shasokush	shusakesh
	sekoshos	sokoshos	shekasish	shusakish
	seshokos	soshakos	shekesosh	shusekish

Appendix B

EXPERIMENT 1: TEST STIMULI

Rules	SH			*SH	
SH	sakesas	sosakas	sheshekash	sakeshes	sakesash
	sakisos	sosakis	sheshokash	sakishos	sakosash
	sakuses	sosakos	sheshokish	sakushes	sekasish
	sasakus	sosikas	sheshukush	sashakus	sekisosh
	sasokos	sukasus	shikeshash	sashekes	sekosesh
	sasukes	sukesas	shikishosh	sashokos	shakishus
	sekeses	sukesus	shikoshish	sashukes	shakushus
	sekisos	sukisas	shishokesh	sekeshes	shekeshas
	sekoses	susekus	shishokish	sekoshos	shekeshus
	sesikos	susikas	shishukosh	seshokos	shekoshos
	sesukis	susokis	shokashesh	seshukis	shikishos
	sesukos	susukos	shokeshish	seshukos	shikoshis
	sikasos	shakashosh	shokushish	shakasosh	shikushis
	sikasus	shakishush	shokushosh	shakosash	shokashes
	sikisis	shakushish	shoshokesh	shakusish	shokushis
	sikosus	shakushush	shoshokush	shasakosh	shukeshos
	sisakus	shashakosh	shoshukish	shasekush	shukoshos
	sisekos	shashekush	shoshukosh	shasokash	sikisash
	sisokos	shashikush	shukeshish	shasokush	sikisish
	sisokus	shashokush	shukoshesh	shekasish	sikosush
	sokasas	shekeshash	shukoshosh	shekesosh	sokasish
	sokasis	shekeshosh	shushakesh	shesakish	sokisosh
	sokosis	shekeshush	shushakish	shesokash	sukasush
	sokusas	shekoshash	shushekish	shesokish	sukisash
	FL			*FL	
FL	sakesas	shukesish	shakashosh	sakesash	sashekesh
	sakeshes	shukoshosh	shakusish	sakosash	sashekish
	sashakus	shusakesh	shasakosh	sekasish	sashekosh
	sasokos	shusakish	shashakosh	sekosesh	sashikush
	sekeses	shushakish	shekeshash	shakeshes	seshekash
	sekeshes	sikashos	shekoshash	shakishus	seshokosh
	seshokos	sikasos	shesakish	shakushus	shasakas
	seshukos	sikushus	sheshekash	shekeshas	shasakis
	sesukos	sisekos	sheshokash	shekeshus	shasekas
	shakasosh	sishokos	shesokash	shekoshos	shasekis
	shakushish	sokoshos	shishokish	shikishos	shasokas

	shashekush	sokosis	shisokesh	shikushis	shesakis
	shasokash	sosakos	shokushosh	shokashes	shesikos
	shasokush	soshakos	shosikosh	shokushis	shisikas
	shekasish	sukasus	shosokesh	shukeshos	shisokus
	shekeshosh	susokis	shusekish	shukoshos	shosakis
	sheshokish	sakuses	sikasus	sikisash	shosikos
	shikeshash	sashokos	sikushis	sikisish	shusikas
	shikisash	sashukes	sisakus	sikosush	sishokesh
	shishokesh	sasukes	sokoshis	sokasish	sishokish
	shisikash	sekisos	sosakis	sokisosh	sishokosh
	shokashesh	sekoshos	sukeshos	sokusash	sishukish
	shokisosh	seshukis	sukisas	sukasush	soshukish
	shoshokush	sesikos	sushukos	sukisash	sushekosh
	IFL			*IFL	
IFL	sakeshes	shekasish	sikashos	sakesash	sashekesh
	sakishos	shekesosh	sikashus	sakosash	sashekish
	sakushes	shekosash	sikushas	sekasish	sashekosh
	sashakus	shesakish	sikushis	sekosesh	sashikush
	sashekes	shesokash	sikushos	shakeshes	seshekash
	sashokos	shesokish	sikushus	shakishus	seshokosh
	sashukes	shesukush	sishakus	shakushus	shasakas
	sekeshes	shikesash	sishekos	shekeshas	shasakis
	sekoshos	shikisash	sishokos	shekeshus	shasekas
	seshokos	shisikash	sishukis	shekoshos	shasekis
	seshukis	shisokesh	sokashas	shikishos	shasokas
	seshukos	shisukosh	sokoshas	shikushis	shesakis
	shakasosh	shokesish	sokoshis	shokashes	shesikos
	shakosash	shokisosh	sokoshos	shokushis	shisikas
	shakusish	shokusosh	soshakas	shukeshos	shisokus
	shasakesh	shosikosh	soshakos	shukoshos	shosakis
	shasakish	shosokesh	soshikas	sikisash	shosikos
	shasakosh	shosokush	sukeshas	sikisish	shusikas
	shasekesh	shosukosh	sukeshos	sikosush	sishokesh
	shasekish	shukesish	sukeshus	sokasish	sishokish
	shasekosh	shukosesh	sushekos	sokisosh	sishokosh
	shasekush	shusakesh	sushekus	sokusash	sishukish
	shasokash	shusakish	sushokis	sukasush	soshukish
	shasokush	shusekish	sushukos	sukisash	sushekosh

Appendix C

EXPERIMENT 3: TRAINING and TEST STIMULI

Sibilant Harmony	Training Stimuli		Test Stimuli: Familiar	Test Stimuli: Novel-Fit	Test Stimuli: Novel-Violate
	sakesas	shakeshash	sikasus	sakases	sakoshis
	sakisos	shakoshesh	sokosis	sakisis	sekashos
	sakosis	shakoshush	sokusas	sekosas	sikeshas
	sakuses	shakushosh	sukisis	sekusus	sokashus
	sekasus	shekashash	sukoses	sikusos	sukishos
	sekesus	shekashish	sakosis	sikisos	shakisesh
	sekisas	shekashosh	sakuses	sokesis	shekusish
	sekusos	shekoshush	sekasus	sokases	shikosush
	sikases	shikeshosh	sekisas	sukesus	shokesash
	sikasus	shikishash	sikesos	sukosas	shukusesh
	sikesos	shikishush	shekashash	shikoshash	sakishus
	sikoses	shikoshosh	shikishash	shekushash	sekoshas
	sokasis	shokishash	shakoshesh	shukashesh	sikushos
	sokisus	shokishish	shukushesh	shakishesh	sokeshes
	sokosis	shokushesh	shokushish	shokeshish	sukashis
	sokusas	shokushish	shukashish	shekoshish	shakusush
	sukesas	shukashish	shikoshosh	shukishosh	shekasish
	sukisis	shukeshesh	shakushosh	shikashosh	shikesosh
	sukoses	shukeshush	shikishush	shakeshush	shokisash
	sukusos	shukushesh	shekoshush	shokushush	shukosesh

Appendix D

EXPERIMENT 4 and 5: TRAINING and TEST STIMULI

Condition	Train	Test Stimuli: SH		Test Stimuli: *SH	
Sibilant Harmony *Strictly-Local k=3, Strictly-Piecewise k=2*	sasa	sasa	shasha	sasha	shosa
	sasu	sase	shashe	sashe	shose
	sesi	sasi	shashi	sashi	shosi
	sesu	saso	shasho	sasho	shoso
	shashe	sasu	shashu	sashu	shosu
	shasho	sesa	shesha	sesha	shusa
	shisha	sese	sheshe	seshe	shuse
	shishi	sesi	sheshi	seshi	shusi
	shisho	seso	shesho	sesho	shuso
	shishu	sesu	sheshu	seshu	shusu
	shosha	sisa	shisha	shasa	sisha
	shoshe	sise	shishe	shase	sishe
	shushu	sisi	shishi	shasi	sishi
	sisi	siso	shisho	shaso	sisho
	sose	sisu	shishu	shasu	sishu
	sosi	sosa	shosha	shesa	sosha
	soso	sose	shoshe	shese	soshe
	sosu	sosi	shoshi	shesi	soshi
	suso	soso	shosho	sheso	sosho
	susu	sosu	shoshu	shesu	soshu
		susa	shusha	shisa	susha
		suse	shushe	shise	sushe
		susi	shushi	shisi	sushi
		suso	shusho	shiso	susho
		susu	shushu	shisu	sushu

Appendix E

EXPERIMENT 6: TRAINING and TEST STIMULI

Condition	Train	Test Stimuli: SH		Test Stimuli: *SH	
Sibilant Harmony *Strictly-Local k=5, Strictly-Piecewise k=2*	sakase	sakasa	shakasha	sakasha	shakasa
	sakesi	sakase	shakashe	sakashe	shakase
	sakiso	sakasi	shakashi	sakashi	shakasi
	sakosu	sakaso	shakasho	sakasho	shakaso
	sakusa	sakasu	shakashu	sakashu	shakasu
	sekase	sakesa	shakesha	sakesha	shakesa
	sekesi	sakese	shakeshe	sakeshe	shakese
	sekiso	sakesi	shakeshi	sakeshi	shakesi
	sekosu	sakeso	shakesho	sakesho	shakeso
	sekusa	sakesu	shakeshu	sakeshu	shakesu
	shakashe	sakisa	shakisha	sakisha	shakisa
	shakeshi	sakise	shakishe	sakishe	shakise
	shakisho	sakisi	shakishi	sakishi	shakisi
	shakoshu	sakiso	shakisho	sakisho	shakiso
	shakusha	sakisu	shakishu	sakishu	shakisu
	shekashe	sakosa	shakosha	sakosha	shakosa
	shekeshi	sakose	shakoshe	sakoshe	shakose
	shekisho	sakosi	shakoshi	sakoshi	shakosi
	shekoshu	sakoso	shakosho	sakosho	shakoso
	shekusha	sakosu	shakoshu	sakoshu	shakosu
	shikashe	sakusa	shakusha	sakusha	shakusa
	shikeshi	sakuse	shakushe	sakushe	shakuse
	shikisho	sakusi	shakushi	sakushi	shakusi
	shikoshu	sakuso	shakusho	sakusho	shakuso
	shikusha	sakusu	shakushu	sakushu	shakusu
	shokashe	sekasa	shekasha	sekasha	shekasa
	shokeshi	sekase	shekashe	sekashe	shekase
	shokisho	sekasi	shekashi	sekashi	shekasi
	shokoshu	sekaso	shekasho	sekasho	shekaso
	shokusha	sekasu	shekashu	sekashu	shekasu
	shukashe	sekesa	shekesha	sekesha	shekesa
	shukeshi	sekese	shekeshe	sekeshe	shekese
	shukisho	sekesi	shekeshi	sekeshi	shekesi
	shukoshu	sekeso	shekesho	sekesho	shekeso
	shukusha	sekesu	shekeshu	sekeshu	shekesu
	sikase	sekisa	shekisha	sekisha	shekisa
	sikesi	sekise	shekishe	sekishe	shekise

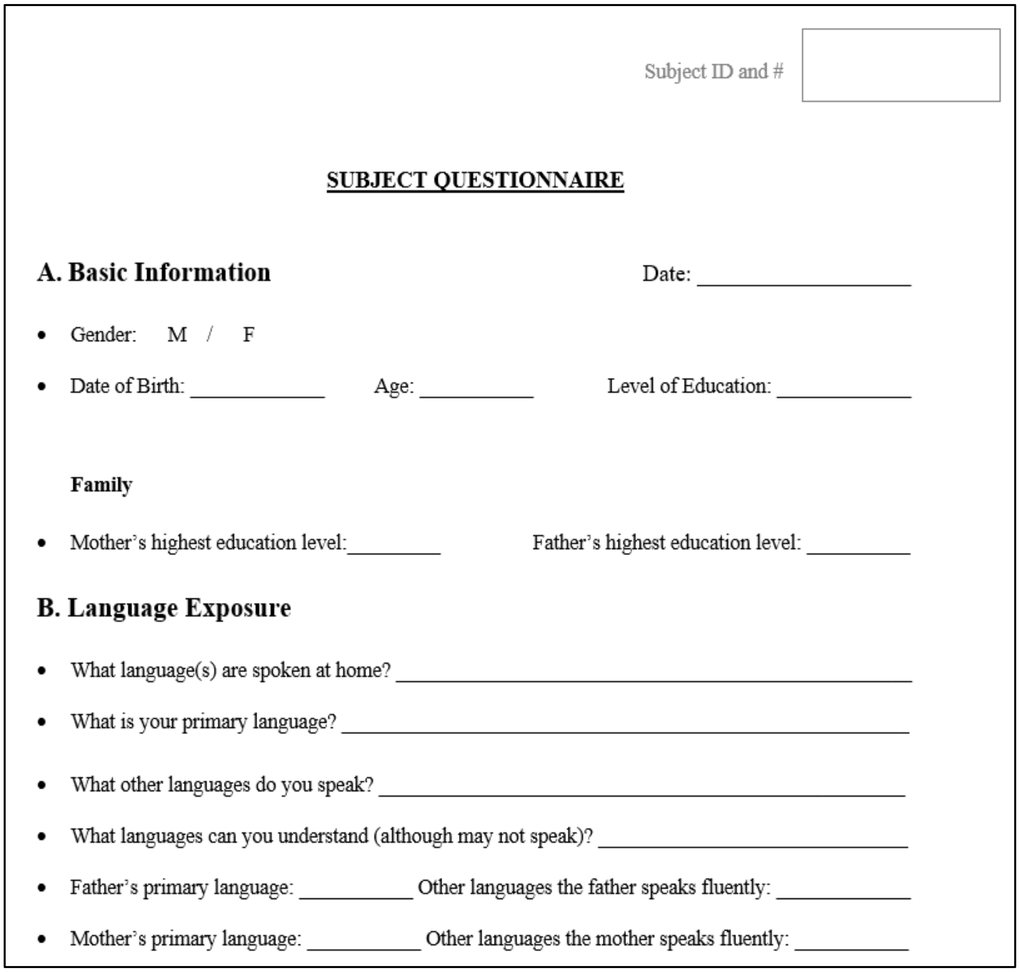
sikiso	sekisi	shekishi	sekishi	shekisi
sikosu	sekiso	shekisho	sekisho	shekiso
sikusa	sekisu	shekishu	sekishu	shekisu
sokase	sekosa	shekosha	sekosha	shekosa
sokesi	sekose	shekoshe	sekoshe	shekose
sokiso	sekosi	shekoshi	sekoshi	shekosi
sokosu	sekoso	shekosho	sekosho	shekoso
sokusa	sekosu	shekoshu	sekoshu	shekosu
sukase	sekusa	shekusha	sekusha	shekusa
sukesi	sekuse	shekushe	sekushe	shekuse
sukiso	sekusi	shekushi	sekushi	shekusi
sukosu	sekuso	shekusho	sekusho	shekuso
sukusa	sekusu	shekushu	sekushu	shekusu
	sikasa	shikasha	sikasha	shikasa
	sikase	shikashe	sikashe	shikase
	sikasi	shikashi	sikashi	shikasi
	sikaso	shikasho	sikasho	shikaso
	sikasu	shikashu	sikashu	shikasu
	sikesa	shikesha	sikesha	shikesa
	sikese	shikeshe	sikeshe	shikese
	sikesi	shikeshi	sikeshi	shikesi
	sikeso	shikesho	sikesho	shikeso
	sikesu	shikeshu	sikeshu	shikesu
	sikisa	shikisha	sikisha	shikisa
	sikise	shikishe	sikishe	shikise
	sikisi	shikishi	sikishi	shikisi
	sikiso	shikisho	sikisho	shikiso
	sikisu	shikishu	sikishu	shikisu
	sikosa	shikosha	sikosha	shikosa
	sikose	shikoshe	sikoshe	shikose
	sikosi	shikoshi	sikoshi	shikosi
	sikoso	shikosho	sikosho	shikoso
	sikosu	shikoshu	sikoshu	shikosu
	sikusa	shikusha	sikusha	shikusa
	sikuse	shikushe	sikushe	shikuse
	sikusi	shikushi	sikushi	shikusi
	sikuso	shikusho	sikusho	shikuso
	sikusu	shikushu	sikushu	shikusu
	sokasa	shokasha	sokasha	shokasa
	sokase	shokashe	sokashe	shokase
	sokasi	shokashi	sokashi	shokasi
	sokaso	shokasho	sokasho	shokaso

	sokasu	shokashu	sokashu	shokasu
	sokesa	shokesha	sokesha	shokesa
	sokese	shokeshe	sokeshe	shokese
	sokesi	shokeshi	sokeshi	shokesi
	sokeso	shokesho	sokesho	shokeso
	sokesu	shokeshu	sokeshu	shokesu
	sokisa	shokisha	sokisha	shokisa
	sokise	shokishe	sokishe	shokise
	sokisi	shokishi	sokishi	shokisi
	sokiso	shokisho	sokisho	shokiso
	sokisu	shokishu	sokishu	shokisu
	sokosa	shokosha	sokosha	shokosa
	sokose	shokoshe	sokoshe	shokose
	sokosi	shokoshi	sokoshi	shokosi
	sokoso	shokosho	sokosho	shokoso
	sokosu	shokoshu	sokoshu	shokosu
	sokusa	shokusha	sokusha	shokusa
	sokuse	shokushe	sokushe	shokuse
	sokusi	shokushi	sokushi	shokusi
	sokuso	shokusho	sokusho	shokuso
	sokusu	shokushu	sokushu	shokusu
	sukasa	shukasha	sukasha	shukasa
	sukase	shukashe	sukashe	shukase
	sukasi	shukashi	sukashi	shukasi
	sukaso	shukasho	sukasho	shukaso
	sukasu	shukashu	sukashu	shukasu
	sukesa	shukesha	sukesha	shukesa
	sukese	shukeshe	sukeshe	shukese
	sukesi	shukeshi	sukeshi	shukesi
	sukeso	shukesho	sukesho	shukeso
	sukesu	shukeshu	sukeshu	shukesu
	sukisa	shukisha	sukisha	shukisa
	sukise	shukishe	sukishe	shukise
	sukisi	shukishi	sukishi	shukisi
	sukiso	shukisho	sukisho	shukiso
	sukisu	shukishu	sukishu	shukisu
	sukosa	shukosha	sukosa	shukosa
	sukose	shukoshe	sukoshe	shukose
	sukosi	shukoshi	sukoshi	shukosi
	sukoso	shukosho	sukosho	shukoso
	sukosu	shukoshu	sukoshu	shukosu
	sukusa	shukusha	sukusha	shukusa

	sukuse	shukushe	sukushe	shukuse
	sukusi	shukushi	sukushi	shukusi
	sukuso	shukusho	sukusho	shukuso
	sukusu	shukushu	sukushu	shukusu

Appendix F

SUBJECT QUESTIONNAIRE



Subject ID and # ☐

SUBJECT QUESTIONNAIRE

A. Basic Information

Date: ____________________

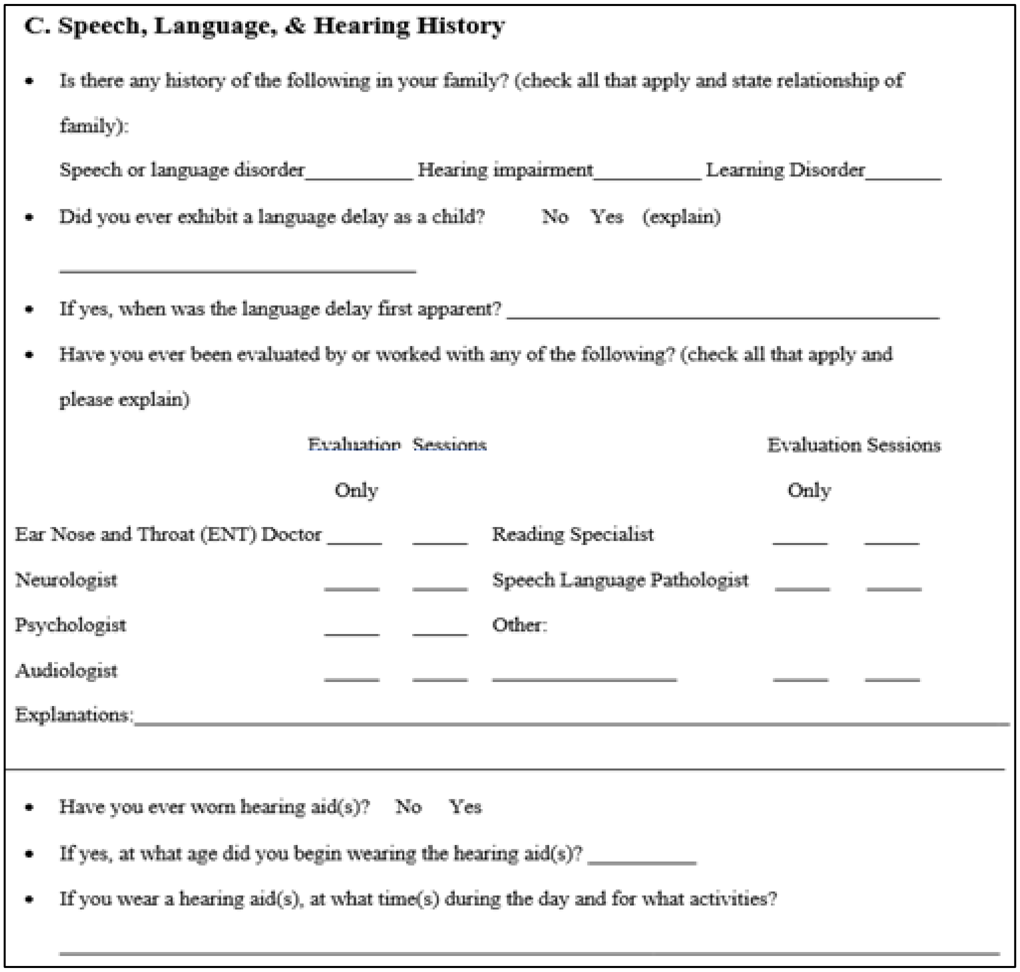
- Gender: M / F
- Date of Birth: _____________ Age: ___________ Level of Education: ____________

Family

- Mother's highest education level:_________ Father's highest education level: __________

B. Language Exposure

- What language(s) are spoken at home? ____
- What is your primary language? ____
- What other languages do you speak? ____
- What languages can you understand (although may not speak)? ______________________________
- Father's primary language: ___________ Other languages the father speaks fluently: ____________
- Mother's primary language: ___________ Other languages the mother speaks fluently: ___________



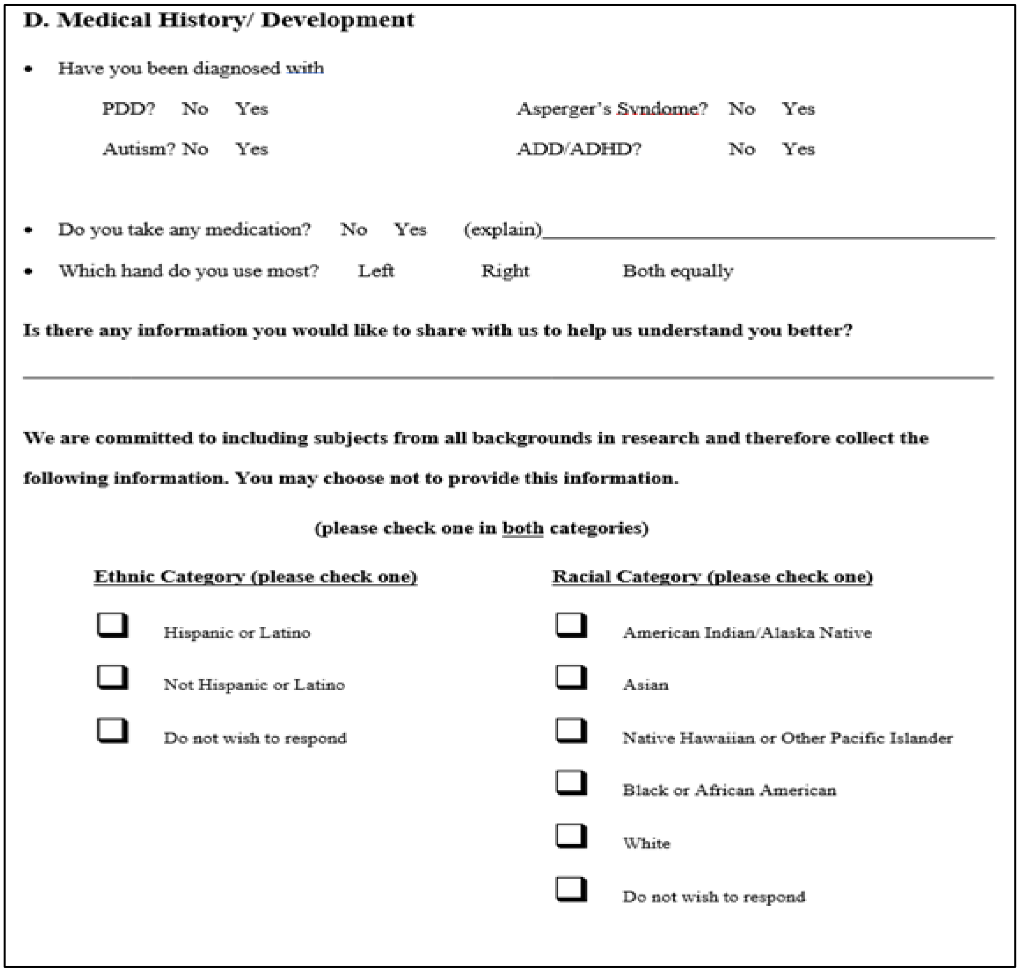
C. Speech, Language, & Hearing History

- Is there any history of the following in your family? (check all that apply and state relationship of family):
  Speech or language disorder__________ Hearing impairment__________ Learning Disorder_______
- Did you ever exhibit a language delay as a child? No Yes (explain)
  ____
- If yes, when was the language delay first apparent? ____
- Have you ever been evaluated by or worked with any of the following? (check all that apply and please explain)

	Evaluation Only	Sessions		Evaluation Only	Sessions
Ear Nose and Throat (ENT) Doctor	_____	_____	Reading Specialist	_____	_____
Neurologist	_____	_____	Speech Language Pathologist	_____	_____
Psychologist	_____	_____	Other:		
Audiologist	_____	_____	_________________	_____	_____

Explanations:____

- Have you ever worn hearing aid(s)? No Yes
- If yes, at what age did you begin wearing the hearing aid(s)? __________
- If you wear a hearing aid(s), at what time(s) during the day and for what activities?
  ____



D. Medical History/ Development

- Have you been diagnosed with

PDD? No Yes — Asperger's Syndome? No Yes

Autism? No Yes — ADD/ADHD? No Yes

- Do you take any medication? No Yes (explain)____
- Which hand do you use most? Left Right Both equally

Is there any information you would like to share with us to help us understand you better?

____

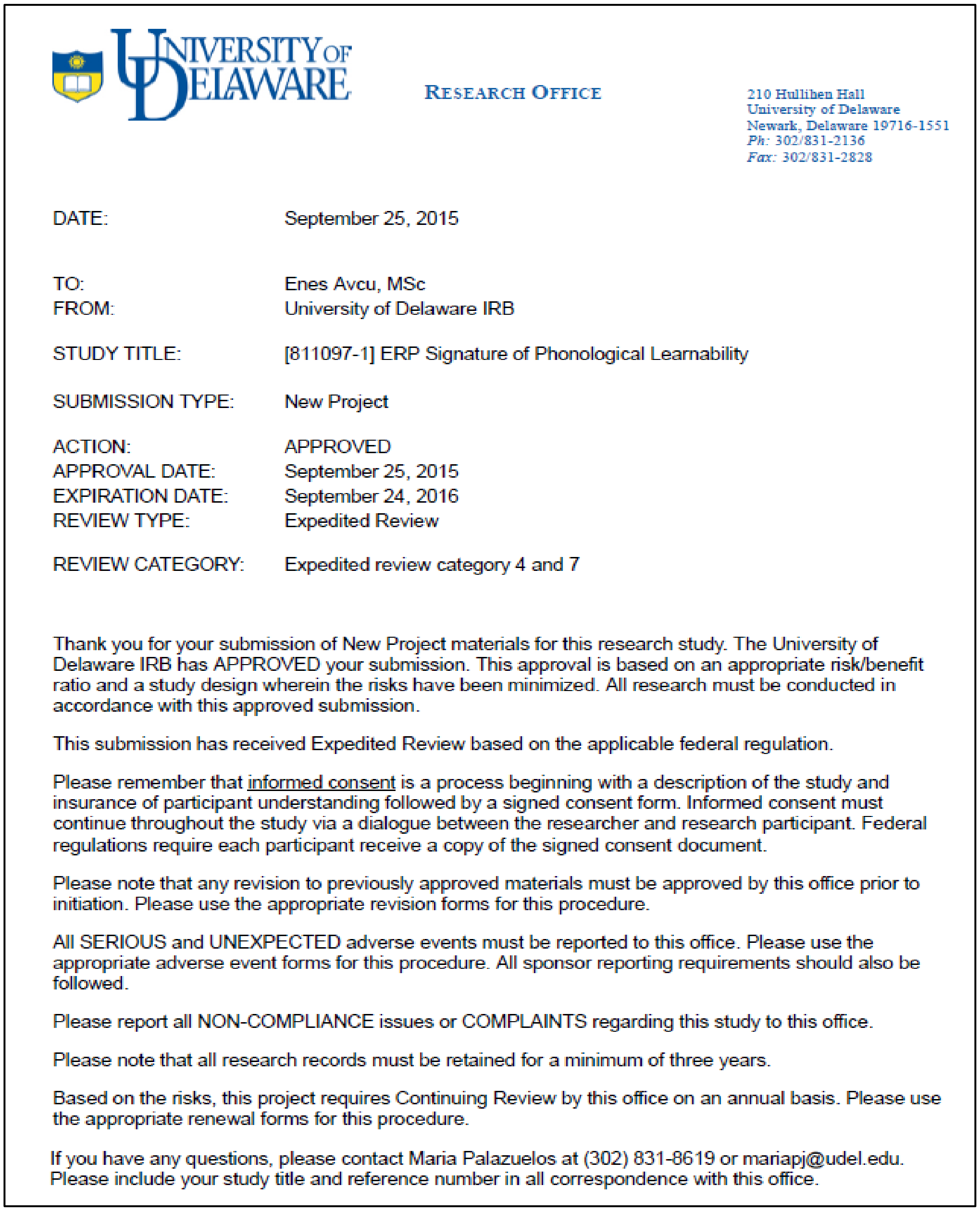
We are committed to including subjects from all backgrounds in research and therefore collect the following information. You may choose not to provide this information.

(please check one in both categories)

Ethnic Category (please check one)	**Racial Category (please check one)**
☐ Hispanic or Latino	☐ American Indian/Alaska Native
☐ Not Hispanic or Latino	☐ Asian
☐ Do not wish to respond	☐ Native Hawaiian or Other Pacific Islander
	☐ Black or African American
	☐ White
	☐ Do not wish to respond

Appendix G

INSTITUTIONAL REVIEW BOARD (IRB) APPROVAL LETTERS



UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE:	September 25, 2015
TO:	Enes Avcu, MSc
FROM:	University of Delaware IRB
STUDY TITLE:	[811097-1] ERP Signature of Phonological Learnability
SUBMISSION TYPE:	New Project
ACTION:	APPROVED
APPROVAL DATE:	September 25, 2015
EXPIRATION DATE:	September 24, 2016
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category 4 and 7

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

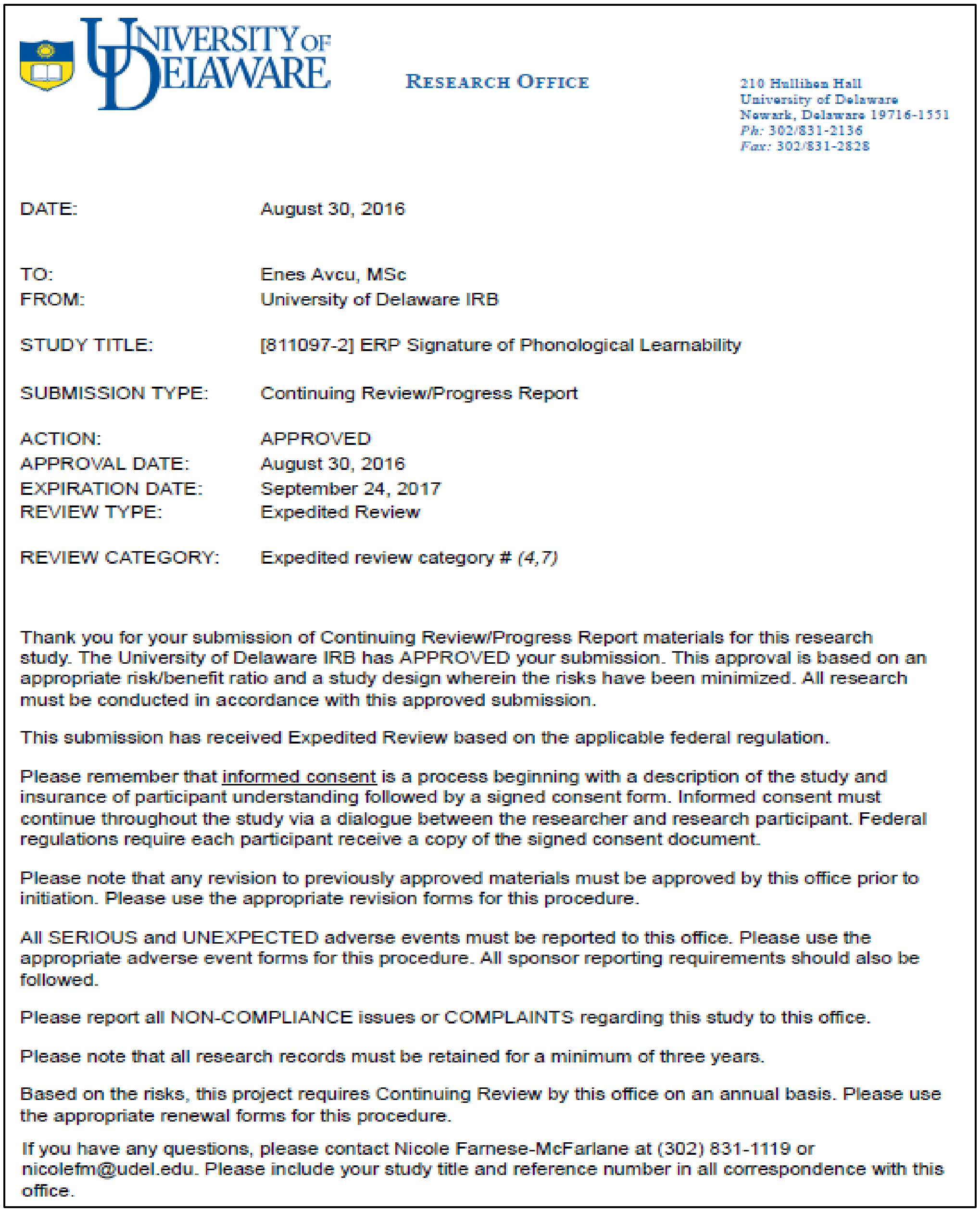
All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Maria Palazuelos at (302) 831-8619 or mariapj@udel.edu. Please include your study title and reference number in all correspondence with this office.



UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: August 30, 2016

TO: Enes Avcu, MSc
FROM: University of Delaware IRB

STUDY TITLE: [811097-2] ERP Signature of Phonological Learnability

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED
APPROVAL DATE: August 30, 2016
EXPIRATION DATE: September 24, 2017
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(4,7)*

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

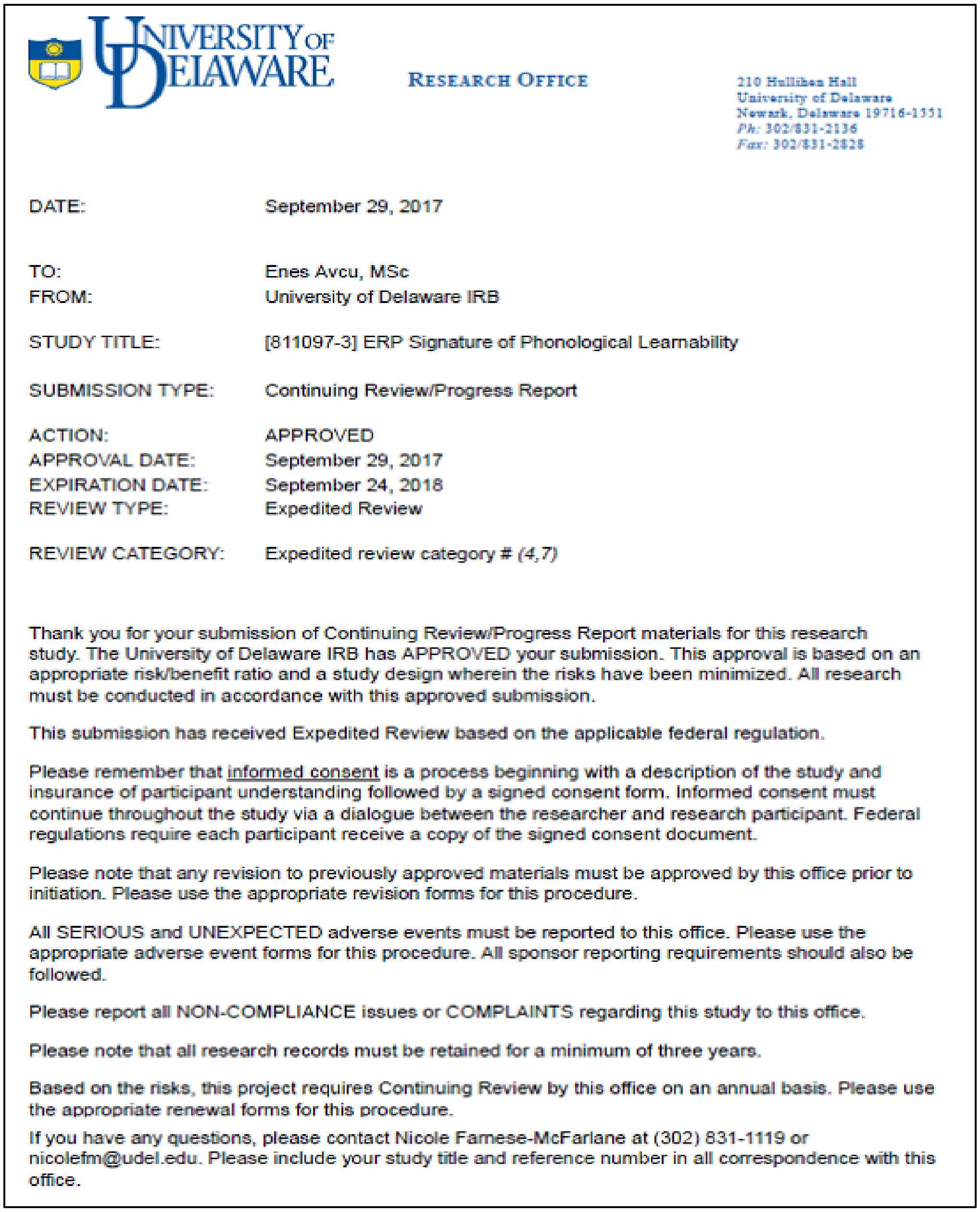
All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.



UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: September 29, 2017

TO: Enes Avcu, MSc
FROM: University of Delaware IRB

STUDY TITLE: [811097-3] ERP Signature of Phonological Learnability

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED
APPROVAL DATE: September 29, 2017
EXPIRATION DATE: September 24, 2018
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(4,7)*

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

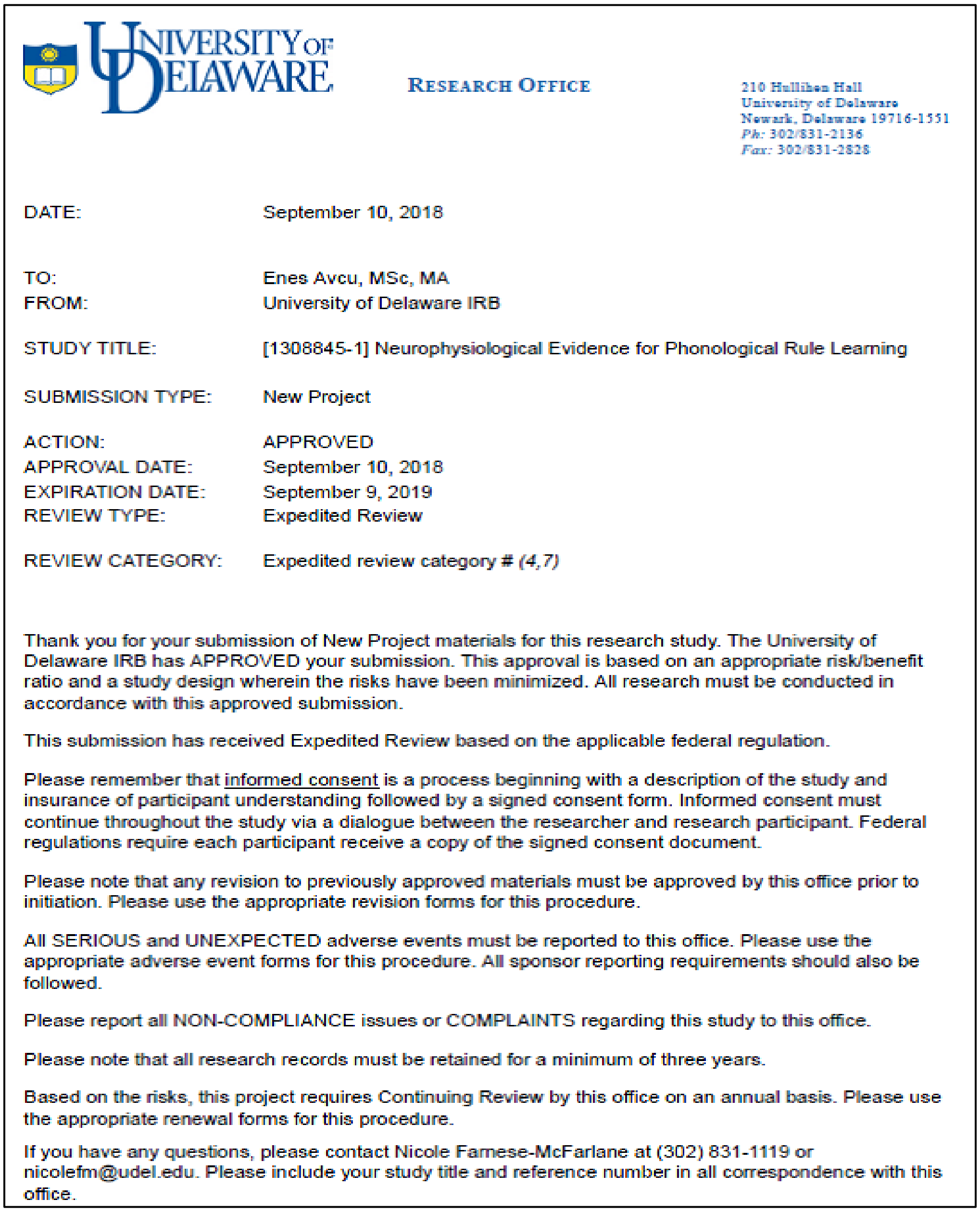
All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.



UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: September 10, 2018

TO: Enes Avcu, MSc, MA
FROM: University of Delaware IRB

STUDY TITLE: [1308845-1] Neurophysiological Evidence for Phonological Rule Learning

SUBMISSION TYPE: New Project

ACTION: APPROVED
APPROVAL DATE: September 10, 2018
EXPIRATION DATE: September 9, 2019
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(4,7)*

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Nicole Farnese-McFarlane at (302) 831-1119 or nicolefm@udel.edu. Please include your study title and reference number in all correspondence with this office.