Disentangling Neural Indices of Implicit vs. Explicit Morphosyntax Processing in an Artificial Language

BY

DAVID ABUGABER

A.B., Princeton University, 2014

M.Phil., University of Cambridge, 2015

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Hispanic Studies in the Graduate College of the University of Illinois – Chicago, 2022.

Chicago, Illinois

Defense Committee:

Laura Batterink, Western University

Jennifer Cabrelli

Philip Hamrick, Kent State University

Kara Morgan-Short, Chair and Advisor

Kimberly Potowski

This thesis is dedicated to Jinyoung Bae.

ACKNOWLEDGMENTS

Thank you to my warm, patient, and supportive advisor Dr. Kara Morgan-Short. Thank you to Dr. Laura Batterink for her generosity in sharing materials and ideas relevant to this project as well as to the other committee members Dr. Jen Cabrelli, Dr. Phillip Hamrick, and Dr. Kim Potowski. Thank you to my colleagues in the Cognition of Second Language Acquisition Lab Dr. Irene Finestrat, Dr. Alicia Luque, Josh Buffington, Marina Ridchenko, and Alexis Berles. Thank you to the research assistants who contributed to data collection, particularly Krishna Bhatia, Will Camacho, Hannah Chaddah, Gayatri Chavan, Brittany Finch, Sarah Hassan, Amber Lewis, and Victoria Ogunniyi. This project was funded by a Dingwall Foundation Fellowship in the Neural Foundations of Language; a National Science Foundation Doctoral Dissertation Improvement Award (#1941189); an American Philosophical Society John Hope Franklin Fellowship, and a Travel and Research Grant from the Institute for Citizens & Scholars (formerly Woodrow Wilson National Fellowship Foundation).

-DA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Summary

Learning new languages is a complex task involving both explicit and implicit processes (i.e., that do/do not involve awareness). Understanding how these processes interact is essential to a full account of second language (L2) learning, but accounts vary as to whether explicit processes help (e.g., DeKeyser, 2007), hinder (e.g., Ellis & Sagarra, 2010), or have no effect on (e.g., Paradis, 2009) the successful acquisition of implicit processing routines. Studies using an artificial language paradigm have established that participants can learn L2 morphosyntactic regularities that they are unaware of (Leung & Williams, 2011, 2012), and a subsequent electroencephalography (EEG) experiment (Batterink et al., 2014) reported different distinct event-related potentials (ERPs) in participants with vs. without awareness of the covert regularity. However, the univariate nature of ERPs makes it impossible to determine whether/to what extent implicit processing occurred in rule-aware learners. Our study addresses this limitation using multivariate pattern analysis (MVPA; Fahrenfort et al., 2018) by training a decoder to detect neural indices of grammar processing at times in the experiment after behavioral measures indicated rule-learning but before participants became rule-aware, and subsequently testing this decoder after participants became rule-aware. We also conduct two follow-up analyses that shed light on the interplay between implicit and explicit grammar processing, asking whether EEG indices of semantic prediction vary between implicit vs. explicit learning as assessed by MVPA, and whether the timing of grammar processing at the neural level is correlated (and thus closely coupled) with behavioral response times.

Following Batterink et al., 52 participants performed a word-classification task that covertly tests for grammar learning by comparing responses to words that follow vs. violate an underlying pattern. Rule-awareness was assessed via systematic debriefing halfway through the

task, at which point the rule was revealed and a final block of trials was performed. Slower reaction times and lower accuracies for rule-violating trials indicated learning even in rule-unaware participants, replicating Batterink et al. However, we did not replicate Batterink et al.'s ERP findings, as we only found a negative deflection in rule-unaware participants and no significant ERP in rule-aware participants. This may be due to natural interindividual variability in ERPs during grammar processing (Tanner, 2019). Furthermore, our MVPA decoding did not show above-chance trial classification accuracy, providing no evidence for the co-occurrence of implicit processing during periods of rule awareness. We also found no evidence for semantic prediction at the neural level in either rule-aware or rule-unaware learners using MVPA. However, for both of these results, follow-up analyses suggested limited decoding sensitivity with the MVPA method on our data in the first place, which was not improved when using a host of alternate analysis parameters. Our ERP-to-reaction time correlation analyses showed evidence in favor of time locking between neural indices of grammar processing and behavioral responses, suggesting a link between the two; however, these results were tempered by the weak ERP effects we found. Overall, the results show strong behavioral learning effects but limited EEG effects. This, along with follow-up behavioral analyses, leads me to question the extent to which learning in this experiment is linguistic vs. non-linguistic. To the extent that our observed learning is linguistic, our results favor weak/no interface models in that rule-unaware and rule-aware participants showed equivalent behavioral performance and there was no MVPA evidence for implicit processing during rule awareness. However, qualitative inspection of the behavioral and debriefing data suggests possible downsides to explicit awareness on task performance. More broadly, this study demonstrates how alternate analysis methods may inform future research on the implicitness/explicitness of second language grammar learning.

CHAPTER I. OVERVIEW

1.1 Introduction

Anybody who has ever tried learning a foreign language in a typical classroom setting might recall memorizing verb conjugation tables, reading long descriptions of when to use and not use certain prepositions, and repeatedly drilling vocabulary lists so as to memorize them as efficiently as possible. At the same time, anybody who has ever been immersed in a foreign language context might have found themselves picking up common words and patterns without much overt effort, almost as if by osmosis. Which, then, might be the better way to learn a foreign language—through approaches that are aware, attentive, conscious, and explicit or unaware, inattentive, subconscious and implicit? Any individual language educator's viewpoint on this fundamental question might radically affect the sorts of pedagogical approaches that are taken. One might imagine a spectrum with explicit metalinguistic instruction on one end and sheer naturalistic immersion on the other.

The field of second language (L2) acquisition has grappled with this question on a theoretical level for several decades now (e.g., Krashen & Terrell, 1983; Doughty & Williams, 1998; DeKeyser, 1998; Hulstijn & De Graaff, 1994; Hulstijn & R. Ellis, 2005), particularly in the past fifteen years or so following the publication of a seminal study which reported that experiment participants showed signs of learning an L2 morphosyntactic regularity even in the absence of awareness (Williams, 2005). More recently, neurolinguistic research has deepened my insight into this topic by combining such experimental language learning paradigms with powerful new tools like electroencephalography (Batterink et al., 2014). This methodology can overcome certain obstacles that arise when studying implicit vs. explicit language acquisition solely through behavioral methods that rely only on externally observable responses (Morgan-

Short, 2014). However, the analysis techniques that have been available for such neurolinguistic data are not equipped to separately measure neural activity tied to implicit and explicit processing in instances when these might co-occur. Disentangling these two neural indices could provide a better window into the underlying cognitive processes involved in processing L2 morphosyntax with and without awareness of a regularity, thus contributing to outstanding debates on the interplay of implicit vs. explicit second language learning.

Thankfully, advances in machine learning approaches from the past few years have increased the amount of insight that can be gleaned from such neural data, allowing us to overcome this obstacle in analysis methods. However, the application of such methodologies to the study of language processing—and particularly *second* language processing—is only in its infancy. This dissertation aims to fill this extant gap in the research by applying innovative analysis techniques to disentangle neural indices of implicit vs. explicit second language morphosyntax processing. In addition to contributing to theoretical debates on implicitness/explicitness in L2 psycholinguistics, the results of this study might inform approaches to second language teaching pedagogy.

1.2 Definitions

Although concepts such as *attention*, *awareness*, *implicitness*, and *explicitness* are critical for this area of research, the exact nature of the constructs denoted by these terms may become blurred at times. To illustrate the difficulty of defining one such term, McLaughlin wrote in 1990 that the word consciousness "[had] acquired too much surplus meaning" to be useful to the field of second language acquisition (1990, p. 617). As such, a more precise definition of the relevant constructs is important. This section aims to define certain key terms that recur throughout this dissertation as they are conceptualized for the purposes of the study. This may not necessarily reflect a clear-cut “correct” definition, and there might not be such a correct definition to the extent that there is some variability in the way that previous researchers have used these terms (Simard & Wong, 2001; Combs, 2004). Nevertheless, the definitions given here align generally with the use of these terms in previous L2 models touching on the topic of implicitness/ explicitness (Robinson, 1995, 2003; Schmidt, 1990, 1995, 2001; Tomlin & Villa, 1994; Leow, 2015) as well as the second language acquisition (SLA) literature overall (e.g., Hulstijn & Ellis, 2005; Sanz & Leow, 2011).

1.2.1 Attention vs. awareness.

As used in this dissertation, the term *attention* refers to a general focusing or tuning of cognitive operations and mechanisms on a particular aspect of the learner’s environment or sensory experience, following the general use of the term in the fields of psychology (e.g., Cowan, 1998; Posner & Petersen, 1990; Delorme et al., 2007) as well as L2 acquisition (e.g., Simard & Wong, 2001; Godfroid et al., 2013). Note that in both the cognitive psychology literature (e.g., Posner & Petersen, 1990) as well as in SLA (as discussed more specifically in Section 2.2.1; Tomlin & Villa, 2014), the concept of attention has been further elaborated with descriptions of subroutines or components such as general alertness, initial orienting to sensory events, and target detection. However, as used in this dissertation, the term “attention” encompasses these components in a more general sense.

More importantly for framing the central topic of this dissertation, attention of itself does not necessarily entail *awareness*, which instead refers to whether information about a given form is consciously registered at a higher conceptual level, beyond mere processing of lower-level features of the input (Schmidt, 1995; Leow, 2015). Just like attention can be analyzed at finer-grained levels, one might define different levels of awareness: Schmidt (1995) distinguishes between *awareness at the level of noticing*—which entails the "conscious registration of the [mere] occurrence of some event" (1995, p. 29) and involves surface level phenomena or item-level learning—and *awareness at the level of understanding*, which refers to "recognition of a general principle, rule or pattern" (1995, p. 29) and involves further abstraction of the linguistic input in terms of systematicities in form-meaning connections. Importantly, though, both of these levels of awareness entail that consciousness regarding the use of a particular form is involved, distinguishing them from the simple lower-level focusing captured by the term attention.

1.2.2 Knowledge vs. processing vs. conditions

The terms *knowledge, processing,* and *conditions* refer to distinct levels at which second language acquisition may be analyzed or manipulated (Hulstijn, 2002; Sanz & Leow, 2011). These might be described as different points in the chain of causation that leads to L2 acquisition, i.e., starting with the external classroom environment and the affordances that it provides, through to individual learners' relatively stable mental representations of the L2 at any point in time, and finally to the actual real-time cognitive mechanisms that invoke said representations during discrete instances of language processing. Each of the levels is defined more precisely below.

In the field of SLA, *knowledge* generally refers to the mental representations tied to a particular structure in one's developing L2 system, without any reference to how these representations are actually accessed and implemented in real time (Hulstijn, 2002; Sanz & Leow, 2011). For instance, L2 knowledge may involve an individual's learned rules and/or patterns for L2 linguistic structure as well as the set of possible utterances that could be produced by those rules, abstracted away from the intervening mental operations involved. In this way, knowledge can be conceptualized as the product of learning, divorced from the processes that created it.

By contrast, *processing* typically refers to the methods of storage of (or means of access to) said rules or regularities (Hulstijn, 2002; Sanz & Leow, 2011). As such, processing emphasizes the actual mechanisms that come into play during actual real-time language perception and production. The distinction between knowledge and processing is most clearly illustrated in cases where a learner shows imperfect or nonstandard proficiency in the L2 even for target forms that they “know” and could successfully perceive or produce under less cognitively demanding (or “more forgiving”) circumstances. In this way, the distinction between knowledge and processing is parallel to the distinction between competence and performance as described by Chomsky (1965), as well as the distinction between the products and processes of second language learning as described by Schmidt (1994).

Bialystok (1990, p. 637) argues that L2 theoretical approaches or models that emphasize knowledge at the expense of processing may be incomplete or overly restrictive if they underemphasize the role or nature of variability in L2 performance during individual instances of processing. According to Bialystok, this incompleteness or restrictiveness arises because these theoretical approaches interpret any observed variability as providing only an imperfect approximation to the learner’s knowledge system. Additionally, because these approaches conceive of linguistic knowledge as a static system, they are unable to easily account for gradient or incomplete amounts of knowledge. In light of these potential issues with focusing only on knowledge (as conceptualized here), the main focus of this dissertation is on second language acquisition at the level of processing.

Finally, beyond the description of second language use or proficiency at the level of individual learners (as captured by the terms knowledge and processing), learning *conditions* refer to the external circumstances in which learners acquire their second language (Robinson, 1996b; Sans & Leow, 2011; Hamrick & Rebuschat, 2014; Tagarelli et al., 2016). Examples of different conditions include naturalistic immersion in an L2 environment and overt L2 instruction in a classroom setting. Critically, this level of description refers only to the external circumstances or environment rather than to learner-internal representations or mechanism of any sort. In this way, individuals within the same learning condition may implement different kinds of processing or acquire different levels of knowledge (Sanz & Leow, 2011).

Note that for brevity in framing the general topic of this dissertation, the terms "learning" and "development" are used here interchangeably to refer to successful L2 acquisition at the level of individual learners more generally, in terms of cumulative changes in processing and knowledge but without necessarily referring to one of these levels in contrast to the other.

1.2.3 Implicitness vs. explicitness

The terms *implicit* and *explicit* may apply separately to either processing, knowledge, or conditions, such that implicit or explicit characteristics might be described for of these levels. In other words, implicit knowledge, explicit knowledge, implicit processing, explicit processing, implicit conditions, and explicit conditions would all refer to distinct (though related) concepts (Hulstijn, 2015), as described below.

In the domain of cognitive psychology, the term *implicit* was first used in regard to learning by Reber (1967) to refer to processes wherein one acquires knowledge about complex, rule-governed input without ensuing awareness of the acquired knowledge. In other words, the knowledge and processing involved in implicit learning are tacit and inaccessible to conscious introspection. Note that, in the SLA context, implicit processing may occur with little implicit knowledge being gained, if changes to the learner's L2 system do not actually occur. Beyond the level of the individual learner, implicit *conditions* would refer to learning circumstances such as naturalistic conversation, information gap tasks, or other activities or settings that do not aim to bring awareness to specific aspects of forms in the L2 input (Norris & Ortega, 2000). As described more fully in Section 2.2.2, however, implicit conditions do not necessarily entail that learners are not also engaging explicit processing/knowledge to cope with task demands.

By contrast, learning is *explicit* when the products or outcomes of learning result primarily in consciously available knowledge (Reber, 1993; Hulstijn, 2002; N. Ellis, 2005; Taylor et al., 2014). However, as noted above, explicit *knowledge* and explicit *processes* technically refer to different constructs. To illustrate this, a learner's willful attempts to discover an underlying form-meaning connection may qualify as *explicit processing*, but if no actual regularity is consciously detected then there would be no ensuing explicit *knowledge* (e.g., if a learner tries but does not successfully invoke any underlying "grammatical rule" or "lexical root"; Leow, 2015, p. 244). In terms of L2 conditions, explicitness is often (but not necessarily) associated with learning conditions wherein individuals are overtly told about underlying patterns or instructed to actively search for such patterns, i.e., when learning is intentional (Rebuschat, 2013). Again, however, explicit conditions may not necessarily entail explicit knowledge or explicit processing. In this way, the dimension of implicitness/explicitness and the dimension of knowledge/processing/conditions may be conceptualized as orthogonal to each other. What ultimately distinguishes explicitness from implicitness, however, is the presence of overt, conscious, and willful mechanisms.

The distinction between implicit and explicit processes in the context of L2 acquisition is corroborated by previous findings from the fields of psychology and cognitive neuroscience. The process of acquiring unconscious knowledge implicitly has been called a fundamental aspect of human cognition (for overviews, see Perruchet, 2008; Reber, 1993; Shanks, 2005), and, beyond comprehension and production of language specifically, implicit processes have been found to underlie many essential skills from non-linguistic domains such as social interaction, music perception, and intuitive decision-making (Berry & Dienes, 1993; Dienes, 2012; Reber, 1993). Further supporting the distinction between implicit and explicit learning, experiments in cognitive psychology have established differences between (relatively more explicit) declarative knowledge and (relatively more implicit) procedural knowledge (Gregory et al., 2016; Knowlton et al., 1996; Stefanucci et al., 2000), and neuroimaging studies have found distinct neuroanatomical bases for these as well (Squire, 2004; Eichenbaum, 2011; Mizumori, 2016). Furthermore, the neuropsychological literature reports that patients with amnesia and other memory impairments (such as the famous case study H.M.) are sometimes able to learn new motor tasks (e.g., drawing complex geometric shapes whose presentation is reversed via a mirror) despite being unable to acquire new declarative information, further supporting a distinction between declarative and procedural learning (Squire, 2009). In this way, empirical studies from non-linguistic domains support the notion of qualitative distinctions between implicit and explicit processing.

In terms of connecting experimental findings from the wider field of cognitive psychology with L2 psycholinguistics specifically, neural components of attention/awareness have also been identified for each of different specific levels of linguistic processing. For instance, for low-level orthographic/word-form processing, one experiment using electroencephalography (EEG, a methodology described in Section 2.2.5) found a triple dissociation between the neural signals tied to either unconscious orthographic processing, visual awareness of seeing a word, or task-relevant responses to words (Schelonka et al., 2017). Similarly, for single word recognition, another EEG study found a difference between early, unconscious perception (indexed via neural oscillations in the gamma range) and conscious perception (indexed by the P300 event-related potential component) (Batterink et al., 2011). At the level of syntax, Batterink and Neville (2013) used an attentional blink paradigm (which involved using a simultaneous tone-categorizing task to distract participants during key parts of experimental trials) to find a difference in event related potentials between consciously vs. unconsciously perceived syntactic violations. In this way, neurolinguistic experiments can reveal differences between conscious and unconscious processing at different levels of language, supporting the notion of implicit and explicit processes as qualitatively distinct.

1.3 Statement of the problem

1.3.1 Is implicit learning of second languages possible?

Models of attention and awareness in the field of second language acquisition disagree as to whether implicit learning—i.e., the acquisition of implicit knowledge through implicit processes—is possible, as argued by Tomlin and Villa (1994) and Leow (2015), or impossible, as argued by Schmidt (1994) and Robinson (2005). This topic has become more prominent following a seminal study that found evidence of L2 morphosyntax acquisition in participants who lack awareness of target forms, suggesting the possibility of implicit learning (Williams, 2005), and the issue received even more attention following failures to replicate in subsequent experiments (Hama & Leow, 2010; Faretta-Stutenberg & Morgan-Short, 2011). More recently, studies using fine-grained reaction-time based methodologies (as opposed to forced-choice task paradigms) have provided additional evidence for the possibility of implicit learning (e.g., Leung & Williams, 2011, 2012), and a recent electroencephalography (EEG) experiment has found differences in the neural activity underlying rule-aware vs. rule-unaware morphosyntax processing (Batterink et al., 2014), suggesting the validity of the rule-aware/rule-unaware distinction in prior behavioral research. However, this represents the only such study (to this author's knowledge) reporting neural differences in participants with different levels of awareness in an L2 morphosyntax learning experiment, making it an important finding to replicate. The first research question addressed in this dissertation is centered precisely on this goal.

1.3.2 What is the relationship between implicit and explicit language learning?

Even if it were clear that implicit learning and explicit learning are separate possible routes to L2 acquisition, the exact interplay between these two would remain to be established. For instance, one might ask whether rule-aware learners' processing also involves to some degree the same kind of processing as rule-unaware learners, with additional, superimposed awareness-related activity. Alternately, it is possible that rule-aware processing "blocks" the kind of learning that occurs in rule-unaware learners. Neurolinguistics research from the past few years has been unable to address these open questions because traditional univariate EEG analysis cannot easily distinguish between separate patterns of neural activity when these occur simultaneously. By contrast, novel multivariate analysis methods that leverage correlations between adjacent data points can allow us to more cleanly disentangle distinct sources of temporally co-occurring neural activity. The second research question for this dissertation uses these new analysis methods to attempt to separately quantify neural indices of implicit vs. explicit processing, allowing us to make inferences about the relationship between the two.

The remaining research questions speak to the interplay of implicit vs. explicit processing in L2 acquisition in a more indirect manner, by further examining the nature of these kinds of processing. The third research question aims to examine whether and to what extent implicit vs. explicit language processing are differentially characterized by real-time linguistic prediction. This would use multivariate analysis of EEG data to detect neural activation linked to the meanings encoded by L2 morphosyntactic forms. The fourth research question gets at the relative roles of implicitness/explicitness in language use by examining the relationship between response times in a linguistic task and the temporal onset of neural signatures of explicit processing, under the intuition that, if explicit processing plays a role in the production of observable responses, then there should be a tight relationship between them in their exact timing. In this way, the three extensions of Batterink et al. (2014) in this dissertation would shed light on the characteristics of and relationship between implicit and explicit language processing.

1.4 Summary of open questions and experiment design

In sum, it is as yet unclear a) whether findings of neural differences between rule-aware (and thus, more explicit) vs. rule-unaware (and thus, more implicit) L2 morphosyntax processing are reproducible; b) to what extent rule-aware processing co-occurs with rule-unaware processing; c) whether rule-aware and rule-unaware processing differ in terms of linguistic prediction; and d) whether responses in a linguistic task are temporally related to (and thus by extension, driven to some extent by) neural activity associated with rule-aware processing.

This dissertation study reproduces and extends a previous EEG study on implicit vs. explicit L2 morphosyntax acquisition (Batterink et al., 2014). The experiment itself would involve exposing participants to artificial language input with hidden regularities; collecting reaction time and accuracy measures of the learning of said regularities; and subsequently gauging participants' awareness of these regularities. The partial replication component of the dissertation employs traditional behavioral and univariate EEG analyses to compare processing in rule-aware vs. rule-unaware participants. Meanwhile, the extension component employs multivariate and response-locked EEG analyses to examine finer-grained differences in neural markers of implicit vs. explicit language processing.

1.5 Organization and preview of the dissertation

The rest of this dissertation is organized as follows. Chapter 2 presents an overview of the relevant research from applied linguistics and psycholinguistics on implicit vs. explicit language learning to date (Section 2.2). It begins by presenting some of the most prominent models in the field of L2 acquisition that touch on the role of attention/awareness in language acquisition, focusing specifically on how these models differ in their respective predictions of whether acquisition can occur implicitly (Section 2.2.1). Then, it provides an overview of previous empirical studies from the field of L2 acquisition that speak precisely to this question (Section 2.2.2). It goes on to describe a series of studies regarding the role of rule awareness in an artificial language experimental paradigm, the results of which suggest the possibility of implicit L2 acquisition (Section 2.2.3). The chapter then lays out some of the issues that arise when using entirely behavioral paradigms to study learning without awareness of underlying rules (Section 2.2.4) and then goes on to describe how these issues can be overcome using electroencephalography (EEG) (Section 2.2.5). Finally, the chapter zooms in on the results of a particular EEG study examining implicit L2 morphosyntax acquisition (Batterink et al., 2014) before motivating a replication of said study (Section 2.2.6).

Going beyond the question of whether implicit L2 morphosyntax learning is possible in the first place, the second half of Chapter 2 discusses what the possible relationship between implicit and explicit processing might be (Section 2.3). It begins by outlining various competing models in L2 acquisition that propose either a strong interface, a weak interface, or no interface between implicit and explicit language learning (Section 2.3.1). It goes on to describe several previous empirical studies in L2 acquisition that are informative to this interface debate (Section 2.3.2). The chapter then describes several insights regarding the relationship between implicit vs. explicit learning that can be gleaned from prior non-linguistic experiments from the wider field of cognitive psychology (Section 2.3.3), before presenting several caveats against generalizing findings from non-linguistic domains to linguistic ones (Section 2.3.4). This section ends with a call for approaches that combine the empiricism of experimental methodologies from psychology and neuroscience with the environmental validity of linguistic stimuli and tasks.

Chapter 2 goes on to describe how the previously described EEG study on implicit L2 morphosyntax learning (Batterink et al., 2014) could be extended to inform interface-related debates in psycholinguistics, specifically proposing three extensions (Section 2.3.5). The first extension involves using a novel machine learning methodology called Multivariate Pattern Analysis (MVPA) to disentangle temporally overlapping neural markers of implicit vs. explicit language processing in my EEG data. The second extension involves using this MVPA methodology to examine linguistic prediction for a morphosyntactically encoded meaning (in this case, animacy) as a possible feature that distinguishes explicit from implicit language processing. The third and final extension involves a fine-grained analysis of the temporal relationship between reaction times in a behavioral linguistic task and previously reported EEG indices of explicit processing that would allow us to infer the relationship between externally observable L2 performance and neural markers of conscious processing. The chapter then ends with a recap of the research questions and an overview of the dissertation study (Section 2.4).

Chapter 3 outlines the materials and methods for the dissertation experiment. The first half of the chapter describes the participant population (Section 3.1), artificial language stimuli (Section 3.2), and procedure (Section 3.3) for my EEG study. Section 3.4 presents the analysis approaches that will be taken to address each of the replication and extension components of the dissertation. These include behavioral (i.e., reaction time and accuracy-based) analyses; traditional EEG analyses using the univariate event-related potential (ERP) method; the novel Multivariate Pattern Analysis technique; and fine-grained methods for quantifying ERP component latency and its relationship to reaction times.

Chapter 4 presents my analysis results. Section 4.1 gives the behavioral (i.e., reaction time and accuracy) results, which indicate that participants did indeed learn the covert grammatical regularity in the dissertation experiment. Section 4.2 gives the results of the Event-Related Potential analysis, which does not reproduce Batterink et al.'s (2014) ERP findings in regard to Research Question 1, i.e., determining whether event-related potentials can capture differences in learning of a covert morphosyntactic regularity between rule-aware and rule-unaware learners in my experiment. Contradicting Batterink et al.'s (2014) findings of negative ERP deflections for rule-unaware learners and positive ERP deflections for rule-aware learners, I found only a weak positive deflection in rule-unaware learners. Section 4.3 gives the results of my Multivariate Pattern analysis on rule-adhering vs. rule-violating trials, which extends Batterink et al. (2014) as per Research Question 2 to determine whether rule-aware participants show implicit processing of the rule at a neural level. Although I did not find statistically significant evidence for the co-occurrence of implicit and explicit grammar processing, additional MVPA analyses suggested that my decoding method was not sensitive enough to detect these neural signs in the first place. Section 4.4 gives the results of the Multivariate Pattern analysis testing the artificial language articles in isolation on a decoder trained on animate/inanimate nouns, which extends Batterink et al. (2014) as per Research Question 3 to determine whether rule-aware and rule-unaware participants show neural evidence of semantic prediction from the covert morphosyntactic regularity in my experiment. I did not find significant evidence for semantic prediction; however, follow-up analyses suggested that the MVPA decoding method was not sensitive enough to detect neural signs of processing of the animate/inanimate distinction in the first place. Section 4.5 gives the results of my analyses correlating reaction times with the latency of the P600 component, which extends Batterink et al. (2014) as per Research Question 4 to determine whether event-related potential markers of explicit processing are closely tied to the production of external behavioral responses. I present mixed evidence in that, although ERP image and trial binning analyses suggested time-locking of ERPs to reaction times, Woody filtering and intertrial phase coherence analyses did not show statistically significant evidence for this time locking.

Chapter 5 presents a discussion of my study results in light of pre-existing theory in the field as well as the relevance of my results for future research. Section 5.1 gives a general interpretation of my main study results confirming that my experiment design yielded signs of grammar learning from the task, in that I reproduced the reaction time and accuracy effects reported in prior literature suggesting learning in the absence of awareness. Section 5.1.1 presents a discussion of my first research question, noting that although I did not find the same ERP effects as reported by Batterink et al. (2014), other work in the ERP literature (Tanner, 2019) suggests that individual differences in ERP responses even among native speakers may mean that comparisons of group means may not be warranted in the first place. Section 5.1.2 presents a discussion of my second research question. Although I did not find significant evidence for overlap of implicit and explicit grammar processing at a neural level, low sensitivity of the MVPA decoding method to the target neural signatures of grammar processing may mean that this was a false negative. Second 5.1.3 presents a discussion of my third research question. I found no significant evidence for semantic prediction in either rule-aware or rule-unaware participants. However, again, these findings are tempered by my low sensitivity of the MVPA decoding to neural activity associated with processing of the target semantic meaning. Section 5.1.4 presents a discussion of my fourth research question. I note that, although the evidence for reaction time-to-ERP correlations was mixed in my inferential analyses, a closer look at my data suggests that the weak nature of the ERP signature in my participants (perhaps due to large individual differences in ERPs) may have been the reason for this (cf. stronger ERPs from tasks using grammar violations during native language listening, which may be more salient).

Section 5.2 discusses the extent to which learning in my experiment may have been driven by lower-level, domain-general aspects of the task rather than linguistic processing per se. I present evidence from an assortment of follow-up analyses that suggest non-linguistic mechanisms of learning, including qualitative analyses of participant debriefing responses (Section 5.2.1), drift-diffusion modeling of behavioral data (Section 5.2.2), behavioral results from alternate versions of the experiment that guard against low-level learning (Section 5.2.3), computational semantic analyses of the effects of prototypicality of each stimulus noun's (in)animacy in my data (Section 5.2.4), and time series analyses illustrating how learning varied as a gradient function of the running statistics of my experimental stimuli rather than as binary learning of a categorical rule (5.2.6). However, in Section 5.2.7 I present several arguments for why learning in my experiment was at least partly linguistic, concluding that the experiment featured both linguistic and non-linguistic components of learning (Section 5.2.8). I go on to discuss my study results in light of previous literature on the implicitness and explicitness of L2 learning (Section 5.3), with a particular eye toward Batterink et al. (2014; Section 5. 3.1), toward theories of whether learning grammar in the absence of awareness is possible (Section 5.3.2), and toward prior research on the interface between implicit and explicit processing (Section 5.3.3). Then, I outline several possible caveats and limitations for the interpretations of my study results, with a particular eye to how the relative trade-off between implicit and explicit processes may depend on the specific circumstances of L2 acquisition, e.g., individual learner profiles, the target L2 forms, and the contexts of learning (Section 5.4). Finally, the dissertation concludes by touching on the potential relevance of my findings on the interplay of implicit vs. explicit processing to L2 psycholinguistic theory as well as to second language teaching praxis (Section 5.5).

CHAPTER II. LITERATURE REVIEW AND MOTIVATION

2.1 Introduction

Several prominent theoretical models of second language acquisition (SLA) agree that some minimal level of attention to target forms is necessary for second language (L2) morphosyntax learning to occur, where attention refers to a focusing of lower-level cognitive mechanisms on particular aspects of the linguistic environment or the learner's sensory experiences (as defined in Section 1.2.1). However, these models disagree on the exact role of awareness of the underlying pattern involved (Schmidt, 1990, 1995; Robinson, 1995, 2003; Tomlin & Villa, 1994; Leow, 2015). In the past fifteen years, a seminal line of research has shed light on this theoretical disagreement through experimental paradigms involving a covert animacy rule in an artificial language. These studies' results suggest that learning of L2 morphosyntax can occur without awareness of an underlying regularity in the form-meaning assignments in the stimuli (Williams, 2004, 2005; Leung & Williams, 2012). More recently, an electroencephalography (EEG) study using a similar design (Batterink et al., 2014) extended these findings by showing event-related potential (ERP) components in the EEG data indicating that purportedly rule-aware and rule-unaware participants were indeed processing the stimuli differently at a neural level. To date, however, there is a gap in my understanding regarding the interplay between L2 processing that is implicit (i.e., occurs with no relation to rule awareness) vs. explicit (i.e., involves some degree of rule awareness), particularly in regard to whether these processes work with or against each other. The following section describes the competing models of attention and awareness in the field of SLA in further detail.

2.2 Is learning without awareness possible?

This section explores the question of whether learning L2 morphosyntax without rule awareness is possible. It begins with an overview of prominent models of attention and awareness in L2 before discussing previous SLA experiments that speak to the issue of implicit vs. explicit learning. It then discusses a specific line of studies that have investigated this topic using a controlled artificial language paradigm, which addresses certain methodological limitations from other experimental approaches. The section ends by laying out the motivation for the EEG replication study that forms the foundation for this dissertation.

2.2.1 Models of attention and awareness in SLA

Interest in the topic of attention and awareness in SLA goes back to Krashen's (1977, 1979, 1981, 1994) influential theories of L2 development which outlined a distinction between "acquisition" and "learning." Under his so-called Natural Approach (Krashen & Terrell, 1983), acquisition is defined as the development of L2 competence for "real communication," whereas learning is defined as "knowing about" or "formal knowledge" of a language (p. 26). These definitions parallel current notions of implicitness and explicitness as conceptualized by other L2 researchers, a distinction that has become so important for the field of SLA that a thematic issue of *Studies in Second Language Acquisition* was centered precisely around this topic both in 2005 (Hulstijn & R. Ellis, 2005, Eds.) and again in 2015 (Andringa & Rebuschat, 2015, Eds.). As noted by Andringa and Rebuschat (2015, p. 186), the interplay of implicit and explicit processes in SLA is relevant for understanding topics as diverse as the differences between child and adult L2 acquisition; disparities in acquisition outcomes across different L2 target features; and the role of learning circumstances (e.g., instructional methods, L1 backgrounds, learner cognitive profiles) in ultimate L2 attainment, among others, leading them to write that "the explicit-implicit distinction permeates through just about every major theme in the study of SLA."

The second language acquisition literature as a whole has agreed that a higher amount of directed attention to target linguistic features promotes L2 development (for a review, see Long, 1996), and this has been supported by meta-analyses of previous research on the topic (e.g., Norris & Ortega, 2000; Spada & Tomita, 2010; Goo et al., 2015). This focus on attention as a key factor in L2 acquisition also underlies pedagogical approaches such as focus on form (Doughty, 2003; Doughty & Williams, 1998; Park, 2005); input enhancement (Lee, 2007; Wong, 2003; Leow, Egi, Nuevo, & Tsai, 2003), and processing instruction (Morgan-Short & Wood-Bowden, 2006; Benati, 2004; Fernández, 2008; VanPatten & Oikkenon 1996; VanPatten & Cadierno, 1993; Wong, 2004; Morgan-Short & Wood-Bowden, 2006).

In contrast to attention, there is less agreement in the SLA literature about the exact role of *awareness*, i.e., the conscious and explicit understanding of meaningful patterns or regularities in the input at a more abstract level (as defined in Section 1.2.1). This debate on the role of awareness has antecedents in the fields of philosophy and psychology, with disagreement between researchers who thought that consciousness was an important line of inquiry (Mandler, 1975; Brewer, 1974; Dawson & Schell, 1987; Lewis & Anderson, 1985) versus researchers who thought that unconscious processes were more important than conscious processes (Freud, 1915; Bowers & Meichenbaum, 1984; Kihlstrom, 1984; Chomsky, 1965, 1980, 1986) versus researchers who thought that consciousness was a non-meaningful or non-scientific term altogether (Lyons, 1986; Nisbett & Wilson, 1977). In the field of SLA specifically, some of the earliest researchers to broach this topic posited that conscious understanding (i.e., explicit knowledge) of a target L2 system is necessary for acquisition (e.g., Bialystok, 1978; Rutherford & Sharwood-Smith, 1985). By contrast, other researchers posited that language learning can proceed entirely through unconscious processes (Seliger, 1983; Krashen, 1981, 1983, 1985;

Gregg, 1984). Meanwhile, a third camp of researchers remained totally agnostic to the issue (e.g., McLaughlin et al., 1983; Odlin, 1986). To date, the most prominent models of attention and awareness in the study of L2 acquisition come from Schmidt (1990), Tomlin and Villa (1994), Robinson (1995), and Leow (2015), as described below.

Schmidt's (1990; 1993; 1994;1995; 2001) model of attention in SLA posits that both low-level perception and higher-level awareness are necessary for L2 morphosyntax learning to occur. In a seminal paper presenting this model, Schmidt (1990) defines several discrete levels of awareness, which are based broadly on the work of prior researchers in the wider study of consciousness (Barušs 1987; Battista 1978; Bowers & Meichenbaum, 1984; James, 1890; Lunzer, 1979; Natsoulis, 1987; Oakley, 1985; O'Keefe, 1985; Tulving, 1985). These three levels of awareness as defined by Schmidt (1990) are:

- *Perception*: this involves the creation of internal representations of external events at the lowest possible level.
- *Noticing*: this level encompasses information that is perceived *and* additionally induces the lowest possible level of subjective experience. In other words, noticing induces a private—i.e., learner internal—experience that is available for later verbalization.
- *Understanding*: this level describes information that is not just perceived and noticed but also analyzed and/or compared to things that have been noticed on other occasions. This implies conscious reflection and an attempt to comprehend the percept's significance, e.g., as in problem-solving and metacognition.

Schmidt's Noticing Hypothesis explicitly states that learning cannot be subliminal. In other words, it cannot proceed through sheer perception alone without any sort of noticing, i.e., awareness of what is being learned (1990, p. 149). Rather, according to Schmidt noticing is the necessary and sufficient condition for input to become intake and thus to be consciously registered and available for further processing (1990). Schmidt notes that learning can be *incidental*, wherein awareness of what is being learned emerges gradually even though the learner is not consciously “paying attention” with the deliberate aim of learning a specific form. Critically, however, some awareness of the learned forms must emerge for L2 acquisition to take place under this model. As Schmidt (1990, p. 142) puts it, "you can't learn a foreign language (or anything else, for that matter) through subliminal perception," implying that conscious attention is the only possible pathway to learning. Note that Schmidt does not clearly distinguish between low-level attention and noticing, writing that “noticing is nearly isomorphic with attention.” (1995, p. 1).

In contrast to Schmidt (1990), Tomlin and Villa’s (1994) model claims that low-level attention to specific forms is necessary *and sufficient* for L2 acquisition, even without any sort of accompanying awareness or conscious noticing. This claim is motivated by findings from an L1 psycholinguistics experiment in which prime words that were presented for as little as 30-80 milliseconds (which falls below the threshold of conscious perception) showed facilitatory effects for the processing of subsequent target items, in the context of a reaction time task wherein participants had to determine whether a word was real or not (Marcel, 1983). According to Tomlin and Villa, this experimental result “shows that people can cognitively process words yet not be aware of those words" (1994, p. 193).

Tomlin and Villa’s (1994) model is based on a three-tiered model of attention from cognitive psychology (Posner & Petersen, 1990) which has been supported by subsequent behavioral and neurocognitive studies (e.g., Posner & Dehaene, 2000; Rothbart & Posner, 2001).

This model defines three discrete levels of attention, each describing a subroutine or component involved in the perception of stimuli. These are as follows:

- maintaining a vigilant/alert state,
- orienting to sensory events, and
- detecting signals for focal (i.e., conscious) processing

Building on this framework from the cognitive psychology literature, Tomlin and Villa (1994) argue for a fine-grained examination of the subprocesses at play when L2 input is processed. Critically, they argue that none of attention's three roles—of alerting for stimuli (i.e., preparing to attend), detecting stimuli (i.e., responding initially to sensory events prior to cognitive registration), and orienting towards stimuli (i.e., committing attentional resources)—actually require awareness (1994, p.193). Under this model, attention to specific forms is necessary *and sufficient* for L2 acquisition, even without accompanying awareness.

In contrast to Tomlin and Villa (1994), Robinson (1995; 2003) posits that awareness is necessary to drive acquisition; in other words, attention by itself is not enough. In this regard, Robinson's model is similar to Schmidt's (1990). However, it differs in that it does not treat attention and awareness as one and the same isomorphic process but instead distinguishes them as qualitatively different constructs. More specifically, according to Robinson's model, input that is detected can briefly enter working memory and consequently activate information that was previously stored in long-term memory (as described in prior psychology literature by Cowan, 1993; 1998). This initial process of detection and information activation may occur automatically and outside of awareness. However, this unaware recognition by itself does not entail learning: for acquisition to actually occur, recognized information needs to become *noticed information*, defined by Robinson (2003, p. 655) as "that subset of detected information that receives focal attention, enters short-term working memory, and is rehearsed" (thus implying awareness as the term is used in this dissertation). In this way, Robinson's model combines the construct of detection from Tomlin and Villa's (1994) model with the construct of noticing from Schmidt's (1990) model.

In a more recent model, Leow (2015) posits that L2 input processing with and without awareness might lead to qualitatively different kinds of processing, each of which may be successful in different circumstances depending on factors like the amount of external input and the amount of cognitive effort and background knowledge on behalf of the learner. The model itself specifically names three stages of L2 processing: input processing, intake processing, and knowledge processing. This model is illustrated in Figure 1.

In the first stage of Leow's model—the input processing stage—, input is initially stored in working memory and turned into intake. The stage itself may proceed at three possible levels. At each of these three levels (and especially at the more "shallow" ones), incoming L2 input may be either stored in working memory or "discarded" if not processed further. These are:

- attended intake, wherein low-level attention (or, as Leow calls it, "peripheral attention") to a form occurs without any accompanying cognitive registration, awareness of the data, or higher-level processing.
- detected intake, wherein some amount of "selective attention" (a term used by Leow, 2015 without further elaboration; p. 242), together with a very low level of processing, is minimally paid to the input. That is, the learner "cognitively took note" of the new information without any level of awareness (Leow, 2015, p. 242). At this point, there is a higher possibility of storage in working memory and further processing when compared to attended intake. However, this may depend on learners'

working memory or whether a higher level of processing/cognitive effort is subsequently allocated.

- noticed intake, wherein linguistic data are attended to and cognitively registered with a low level of awareness. This aligns with Schmidt's idea of "noticing" as described above (1990; 1995; 2001). Here, focal attention is accompanied by a low level of awareness. At this point, there is an even higher potential for input to be stored in working memory and made available for further processing.

After the initial input processing stage, the intake processing stage constitutes the interface between preliminary intake and the developing L2 system. Here, intake can either be processed shallowly with little accompanying cognitive effort (e.g., as non-systematic chunks) or more deeply, with conscious encoding/decoding of linguistic information as well as conceptually driven processing. According to Leow (2015), the intake processing stage often involves a higher level of awareness relative to the input processing stage, which helps to maintain the input in working memory. Furthermore, this input can activate the learner's prior knowledge. With repeated instances of activation of prior knowledge for any given form, there comes a reduction in the necessary level of awareness or depth of processing for the learner to successfully process the L2 data (Leow, 2015). Finally, the knowledge processing stage involves an interface between what is processed/produced by the learner and the developing L2 system.

A key prediction of Leow's (2015) model is that L2 acquisition is possible through fully implicit means. In implicit processing, a low level of processing "may potentially lead to" implicit restructuring, but only if certain prerequisites are met. These prerequisites include having a sufficient quantity of input that occurs in a meaningful context and is followed by sufficient time for the learner to process and internalize the exemplars to have "knowledge available" for subsequent usage (Leow, 2015, p. 244). In contrast to implicit learning, explicit learning (which may involve processes such as hypothesis testing and rule formation) may occur even in cases where the stringent prerequisites mentioned above are not met. However, explicit learning requires more prior knowledge activation, depth of processing, and a high level of awareness for linguistic data to be "explicitly restructured if necessary" and "stored in the grammatical system." In this way, under Leow's (2015) model, implicit processing is more automatic but only occurs in special conditions, whereas explicit processing is more flexible but also more effortful.

On a broad level, the four prominent models of attention in SLA described above differ in whether they predict that learning without awareness is possible. Specifically, Schmidt (1990) and Robinson (1995) argue that attention and awareness are both necessary for L2 acquisition to occur, whereas Tomlin and Villa (1994) and Leow (2015) posit that learning without awareness (i.e., implicit learning, without any explicit component) can also be a successful route to acquisition. The section below discusses previous empirical research in the field of SLA that could be informative to this disparity in the models' predictions.

2.2.2 Previous studies on attention and awareness in SLA

What do previous studies in the field of SLA say about the role of attention and awareness in L2 learning? One approach to this question is to manipulate learning *conditions* in an experimental setting by using either implicit or explicit training approaches (e.g., Sanz & Morgan-Short, 2005; VanPatten & Oikkenon, 1996). Meta-analyses of such research (which aggregate and quantify results from prior experiments in the field) have suggested that implicit training conditions can be effective for producing L2 acquisition gains, although they are not as effective as explicit approaches. Perhaps the most widely cited meta-analysis comparing implicit vs. explicit approaches in L2 acquisition comes from Norris and Ortega (2000), who summarized findings from 49 experimental and quasi-experimental studies conducted between 1980 and 1998. The authors found that explicit conditions (defined as conditions wherein rule explanation comprised part of the instruction or wherein learners were directly asked to attend to particular forms to attempt to arrive at metalinguistic generalizations on their own) were generally more effective at inducing target-oriented L2 gains at post-test and delayed-post-test phases than implicit conditions. More specifically, explicit conditions were associated with a large effect size for learning (d = 1.13), whereas implicit conditions showed more modest (but nevertheless significant) gains, with an effect size of d = 0.54.

Similar findings come from a more recent meta-analysis that was conducted specifically to determine whether the complexity of an English grammatical feature was a factor in the effectiveness of implicit vs. explicit conditions (Spada & Tomita, 2010). This meta-analysis of 41 studies found that, both for grammatical features categorized as either simple or complex (a categorization based on Hulstijn & De Graaff, 1994), explicit treatments led to larger gains than implicit approaches in terms of accuracy at test as well as in controlled knowledge and spontaneous use of the target forms. Again, however, implicit methods were nevertheless found to be effective to some degree for inducing gains in immediate and delayed post-tests.

These findings of a strong learning effect from explicit conditions and a weaker yet confirmed effect from implicit conditions was also corroborated by an even more recent meta-analysis by Goo et al. (2015). This meta-analysis sought to carefully control for the amount or level of instruction for implicit and explicit conditions across the analyzed studies. Using strict criteria that aimed to maintain comparability across implicit and explicit conditions, this meta-analysis ultimately reviewed 34 studies, comprising 11 studies included in the original Norris and Ortega (2000) meta-analysis described above as well as 23 studies published in the time period 1999-2011. In the results of this meta-analysis, Goo et al. (2015) found advantages for explicit instruction (with a large effect size at g = 1.290) over implicit instruction (with a medium effect size at g = 0.774). This coincides with the other meta-analyses described above in reporting that gains from implicit training conditions are possible (though not as large as from explicit conditions)

Although such meta-analyses are an effective way of synthesizing a wide span of research findings from the field, they are not without their limitations. For instance, meta-analyses may aggregate research findings that may not be strictly comparable across studies, with variations in participant characteristics, experimental procedures, or other factors limiting comparability. Furthermore, meta-analyses cannot overcome the so-called "file drawer" problem: they do not account for study results that were not published or otherwise reported, e.g., because they did not achieve statistically significant results. A related issue is that meta-analyses may underestimate the effectiveness of implicit instruction because of cumulative bias in research on the topic to date (as noted by Doughty, 2003). For instance, because implicit treatments generally require longer interventions than explicit treatments (N. Ellis, 2005), the short study durations typical for this research (usually around one hour; Rosa & Leow, 2004) may unduly bias study results in favor of explicit conditions (R. Ellis et al., 2009). Additionally, in order for implicit conditions to show learning, they may require very careful elaboration of training materials as well as outcome measures that are highly sensitive to subtle changes in interlanguage development. Perhaps as a result of such difficulties in designing implicit conditions, less than one third of the of 98 distinct instructional treatments identified by Norris and Ortega (2000) involved methods that were categorized as implicit, and only eight of the 49 synthesized studies involved freely constructed response tasks (a highly implicit condition). Further biasing results against implicit methods, participants in such studies often come in with low levels of proficiency in the L2 (typically from classroom settings), which potentially means that they would benefit more from explicit rather than implicit approaches relative to other participant populations (Morgan-Short et al., 2012). In addition, for many of the meta-analyzed studies, the explicit training conditions may have unintentionally provided more input to participants than implicit conditions (because they provide explicit information that itself includes the target forms in addition to the stimuli that would be provided in the implicit condition), meaning that the total amount of input and time-on-task may not be comparable across conditions (Rosa & Leow, 2004; VanPatten & Oikkenon, 1996).

Perhaps the biggest limitation with the meta-analyses described above is that the individual studies involved only compared implicit vs. explicit training *conditions* as opposed to implicit vs. explicit processing *itself* (an issue mentioned in Section 1.2.2). To illustrate this, purportedly implicit conditions like input floods, input enhancement, recasts, etc. do not preclude that students reflected consciously and explicitly on the input that they were exposed to in the training phase. As such, these meta-analyses may ultimately be more useful for determining the relative advantages of implicit versus explicit *instruction* rather than for determining whether L2 morphosyntax acquisition through fully implicit means is possible.

To address this latter question, one would need a more controlled experimental context that operationalizes the implicitness/explicitness of L2 learning at the level of individual participants. As pointed out by Simard and Wong (2001), a major limitation to this endeavor comes from the "difficulty of finding operational definitions for the attentional functions in SLA" (p. 109). One reason for this difficulty is that, in higher level tasks such as language processing, various components of attention such as alertness, detection, and orienting may be activated simultaneously, making their relative individual contributions hard to disentangle (Posner, personal communication cited in Simard & Wong, 2001, p. 110). As Simard and Wong state bluntly, "designing a task that could adequately examine the isolated effects of alertness and orientation during detection of L2 input seems virtually impossible" (2001, p. 110). They go on to write:

> [The question of learning without awareness] could be empirically tested with an experiment that exposed L2 learners to some kind of L2 input that allows for new form-meaning connections to be made below the threshold of awareness. If I find that subjects who detected this [sic] stimuli below the threshold of awareness (i.e., those who were not aware) perform better on some kind of performance measure than those who were not exposed to the stimuli, and perform as well as subjects who were exposed to the stimuli above the threshold of awareness (i.e., those who were aware), then I will have more concrete evidence that awareness is not necessary for detection of L2 input. However, if I find that subjects who were exposed to the L2 input below the threshold of awareness do not perform better than those subjects who were not exposed to the input at all, then there is a problem with the position that awareness is not necessary for detection and learning to occur in an L2 context. (Simard & Wong, 2001, p. 120).

Thankfully, such a hypothetical methodology has arguably already been developed, as described in the section below.

2.2.3 Rule awareness as a test case for awareness in SLA

One way to analyze the role of implicitness vs. explicitness in L2 learning in a controlled manner might go beyond a description or manipulation of the implicitness/explicitness of the external training conditions and examine instead whether individual participants in a study have or do not have rule awareness, i.e., conscious noticing of the form-meaning connections being acquired. This would more closely address the concern that individual learners might actually be using implicit processes and/or knowledge in explicit conditions and vice versa, thus homing in more specifically on implicitness/explicitness as a central topic of inquiry. An additional advantage for rule awareness as the relevant construct for analyzing implicitness/explicitness in L2 learning is that categorizations such as "rule-aware" and "rule-unaware" demarcate straightforward binary categories that are intuitive and accessible for language teachers in describing their students' levels of awareness in a classroom setting. As an additional benefit, rule awareness is relatively easy to manipulate (e.g., by explicitly explaining a form) and to measure (e.g., by collecting participant reports on levels of awareness) in an experimental context (though as described further below, several issues can arise when assessing awareness; see Hama & Leow, 2010; Rebuschat et al., 2013). Finally, choosing rule (un)awareness as my operationalization of implicit vs. explicit learning allows us to build off of a long line of studies in experimental SLA that have taken this approach (Williams, 2004, 2005; Hama & Leow, 2010; Faretta-Stutenberg & Morgan-Short, 2011; Leung & Williams, 2011, 2012, 2014; Rebuschat et al., 2013; Batterink et al., 2014), as reviewed further below.

Previous studies in the field of SLA have used a variety of methodologies to assess rule awareness. These mainly involve first-person reports of experience that are collected in the form of speech or writing through measures such as: diary entries in which learners reflect on their experience over weeks or months (Schmidt & Frota, 1986; Warden et al., 1995); real-time "think-aloud" protocols in which learners verbally describe their subjective experiences while directly engaged in learning activities (Bowles, 2010; Leow & Morgan-Short, 2004); immediate off-line responses to prompts asking them to recall experiences (Philp, 2003); and somewhat more delayed responses to written questionnaires (Bell & Collins, 2009; Robinson, 1996, 2010) or oral interview questions (Leow, 2000). One important caveat is that awareness may have different effects at different levels of language (e.g., phonology, lexicon, morphosyntax; Mackey et al., 2000; Jiang, 2004). However, for the purposes of this dissertation, the role of awareness is only examined specifically in the context of L2 acquisition of morphosyntax.

An early study investigating the relationship between awareness and L2 morphosyntax comes from Green and Hecht (1992), who found a dissociation between learners' performance on verbal reports vs. sentence correction tasks in naturalistic SLA. Specifically, the researchers found that participants' ability to correct grammatical errors did not entail that they could provide accurate explanations for their corrections. Furthermore, accurate corrections were often associated with incorrect explanations. As such, the authors argued that performance on the sentence correction task was driven by implicit rather than explicit knowledge, and thus that awareness by itself was not the sole determiner of L2 morphosyntax performance. However, one possible limitation with this study comes from the open-ended nature of the verbalization task as well as the range of grammatical rules involved, which might not have detected explicit knowledge that could not be eloquently verbalized by the participants.

A more controlled study that examined the role of awareness in the domain of L2 morphosyntax comes from Leow (2000), who assessed the relationship between explicit processes and learners' recognition and production of written morphology (namely, irregular third person forms for stem-changing *-ir* verbs in Spanish) by asking participants completing a crossword-style task to verbally report their concurrent thoughts. Leow found that only learners whose verbal reports indicated awareness of the underlying form showed significant improvement between pre- and post-test phases, suggesting that awareness is important for L2 morphosyntax acquisition. However, it is possible that participants' attempts to verbalize their experience may have distracted attention from the learning activities themselves, negatively affecting performance. Conversely, the learning task might have negatively impaired participants' ability to accurately describe their subjective experiences in real time. Such so-called reactivity effects have been found in one study (Sachs & Polio, 2007) which reported that a think-aloud condition led to diminished processing of feedback, as measured by the percentage of corrections on essay drafts after learners received reformulations of non-targetlike structures (see Bowles, 2010 for extensive discussion of this issue of reactivity).

One possible way to study the effects of awareness on L2 morphosyntax acquisition while avoiding reactivity effects is to collect awareness assessments only during a final experimental phase, subsequent to a main learning task. This was the method taken in a seminal study that sought to examine L2 morphosyntactic learning without rule awareness (Williams, 2004), a study which subsequently inspired a series of replications and extensions (Williams, 2005; Faretta-Stutenberg & Morgan-Short, 2011; Hama & Leow, 2010; Leung & Williams, 2012; Rebuschat et al., 2013). In this original study, 37 participants (described as "predominantly Cambridge University students and researchers with various language backgrounds"; Williams, 2004, p. 217) performed a mini-artificial language learning task wherein they were led to notice a relevant form and its corresponding meaning, but without requiring them to necessarily notice the relationship between the two in a conscious and explicit way. The language itself comprised eight novel nouns learned as translation equivalents of common English words (e.g., *johombe/i* "monkey/s", *nawase/i* "vase/s") and eight articles that were learned as translation equivalents of English "the" (*ig, i, ga, ge*), "a" (*ul, ula*), and "some" (*tei, tegge*). Participants were not informed why there were four words for "the" and two each for "a" and "some." During training, participants were presented with noun phrases consisting of combinations of articles and nouns (e.g., *i johombi, ga nawase, ul johombe, ge nawasi*). For each trial, they had to repeat the noun phrase aloud (to facilitate memory encoding), indicate whether the noun referred to a living or nonliving thing, and translate the phrase into English. This was designed to draw participants' attention to the animacy of the noun without explicitly revealing that this was relevant for the choice of article. After a training period, participants were asked to choose the corresponding translation of short English phrases (e.g., “the monkey”) in the artificial language, in the context of a two-alternative forced choice task in which they chose between two possible noun phrases (e.g., *ig johombe / ga johombe*) wherein both articles were of the correct number and definiteness but one violated the animacy rule. In a final phase of the study, the participants were interviewed to ascertain their level of awareness of the system and whether they attempted to detect a pattern.[1] Of the 37 total participants, 30 did not show explicit awareness of the relationship between determiner choice and accuracy in the post-test questionnaire. However, despite this apparent lack of rule-awareness, these rule-unaware participants showed signs of learning of the animacy regularity, as indexed by above-chance performance in the two-alternative forced choice task. By contrast, a subsequent control experiment established that any assignment of certain prefixes to animacy did not occur when the systematic relationship with animacy was removed, indicating that it was exposure to the input rather than pre-existing biases that produced the apparent learning effect. These results from Williams’ (2004) first experiment suggest that it is possible to learn morphosyntactic form-meaning connections in a second language even when these are not explicitly noticed.

In the results of this first experiment, Williams (2004) found that native proficiency in a language that encodes grammatical gender was a significant predictor of whether or not a

[1] Williams (2004) does not report the exact questions used in the post-experiment interview. He notes only that “the remaining 30 participants said they did not try to work out the system during training, and they were still unaware of the relevance of animacy at the end of the test phase. Two of them seemed to have tried to work out a system during the test phase, but even then only in the latter part (trained items), and they made no reference to animacy” (p. 220).

participant became rule-aware.[2] Thus, a second experiment (also reported in Williams, 2004) was performed using the same design but with a more simplified artificial language so as to reduce the potential impact of this previous language knowledge. In this second experiment, 17 participants (again, university students of various language backgrounds) were taught four novel articles (*gi, ul, ro*, and *ne*) and told that these encoded the distance of a co-occurring English noun, such that two of the articles are used with distant referents and two are used with nearby referents. However, there was also an underlying, untaught regularity in the semantic features encoded by these articles: namely, two of the articles were only ever used with animate nouns, and two articles were only ever used with inanimate nouns. In this second experiment, 14 of 17 participants showed no rule awareness by the end of testing, and these participants also did not show above-chance performance on previously unencountered items during a test phase involving a two-alternative forced choice task. Grammatical gender in the participant's native language was not found to affect performance in this experiment, perhaps due to the more simplified nature of the artificial language. This lack of implicit learning effects in the second experiment (contrary to the results of the first experiment) was attributed by the author to the small training set size (Williams, 2004).

In a follow-up study with a larger sample size that also used the same four novel articles underlyingly encoding animacy, Williams (2005) reported evidence of implicit learning in rule-unaware participants, supporting the findings of the first but not the second experiment in Williams (2004). In this 2005 experiment, the training task was altered so that the participants (described as 41 speakers of mixed language backgrounds, most of whom were university

[2] By contrast, L2 learning experience and phonological short-term memory did not show a relationship with rule awareness.

students) read stimulus sentences, indicated whether a referent was near or far (based on a pseudoword article), and were then asked to mentally imagine the situation described in the sentence, e.g., *The little boy patted* gi tiger *in the zoo* or I *all admired* ne pictures *from the other side of the gallery.* The test phase consisted of a forced-choice task that involved using the novel articles to fill blanks in context sentences (e.g., *The lady spent many hours sewing…*). To assess rule awareness, a structured interview was performed subsequent to the experiment phase wherein participants were first asked to report the criteria that they had used to make their choices; "any references to living or nonliving, moves or does-not-move, and so forth were interpreted as evidence of awareness of animacy" (Williams, 2005, p. 283). Participants who did not report such awareness were told that there was indeed an underlying rule, and then proceeded to perform a second test phase during which they were invited to work out the rule. If by the end of the experiment, subsequent to the second phase, they still did not report the correct pattern, then the rule was overtly explained to them and they were then asked if they had considered animacy as a relevant factor at any point during the experiment. After the first test phase, 8 of the 41 participants identified animacy as a relevant factor. Of the 33 remaining participants, 6 described an incorrect rule (e.g., tied to irrelevant factors such as syntactic structure, phonology, the nature of the verb, etc.), and 27 participants attributed their responses only to familiarity or intuition. Of these 33 initially unaware participants, 11 became aware of the regularity following a second test phase. None of these 33 participants claimed to have considered animacy as the relevant factor during the first test phase, after full debriefing at the end of the experiment. As in the second experiment of Williams (2004), native proficiency in a language with grammatical gender did not affect indices of learning among the rule-unaware learning. However, L2 experience with gendered languages and general experience studying language-related disciplines were each tied to significantly better performance. Most critically for the central topic of this dissertation, above-chance accuracy was shown during the test phase even in participants who reported no awareness of the underlying rule. Similar results were found in a second experiment performed with 24 advanced non-native speakers of English. These findings align with those from the first experiment of the previous study (Williams, 2004) to suggest that L2 morphosyntax learning without awareness is possible.

A subsequent extension study by Hama and Leow (2010) aimed to explore the disparities between Williams' (2004, 2005) findings of implicit learning versus Leow's (2000) aforementioned findings of Spanish morphology learning only in participants whose concurrent verbal reports indicated noticing (and thus, explicit processing) of a target L2 form. In this extension, the researchers adopted the same artificial language paradigm (including a sentence-reading task followed by a two-alternative forced choice task) as in Williams (2005), but with a concurrent think-aloud task. For the sentence-reading task phase, the experimenters asked 96 participants (all individuals without backgrounds in linguistics[3]) to "choose 'near' or 'far' by selecting one of the choices while providing verbal reasoning for their choices by thinking aloud" (Hama & Leow, 2010, p. 475). Based on the productions from this think-aloud task, participants were classified into three groups, based on whether they *noticed* (i.e., mentioned or commented on) some aspect of animacy; *understood* the rules, i.e., explicitly stated the underlying pattern; or did neither. This goes beyond the previous studies' binary categorizations of simply rule-aware vs. rule-unaware, and instead aligns with previous psycholinguistic models' distinctions between lower-level noticing vs. higher-level understanding of linguistic forms (e.g., Schmidt, 1990; Robinson, 1995). Other modifications to Williams's (2005) paradigm are that

[3] No further details are given by Hama and Leow (2010) regarding participant characteristics or recruitment.

four possible choices were given instead of two in the forced-choice task during the final testing phase, and that all tasks were performed in the auditory modality so as to keep modality consistent across the experiment phases (cf. Williams, 2005, who used auditory stimuli at training and the written modality for testing). In contrast to Williams' (2004, 2005) results, Hama and Leow (2010) found no evidence of above-chance performance either on multiple choice or production tasks in the 34 participants who were classified as rule-unaware.[4]

A subsequent replication study that adopted Hama and Leow's (2010) three-tier categorization system of rule awareness aimed to control for differences in native language background while accounting for previous experience with languages that encode grammatical gender (Faretta-Stutenberg & Morgan-Short, 2011). Of the 30 participants in this study (all undergraduate native English speakers recruited through introductory psychology classes at a large university), 15 reported proficiency in an L2, and 11 of these had an L2 that encoded grammatical gender. The same artificial language stimuli, sentence-reading task, and two-alternative forced choice task as in Williams (2005) were employed. The authors found that, although performance on test items was descriptively higher for participants who noticed animacy as a potential factor (without stating the underlying rule) when compared to participants who did not mention animacy at all (as per an extensive structured interview administered at two points in the study[5]), no evidence of learning (i.e., of significant above-chance accuracy on a

---

[4] This sample size of 34 was obtained after extensive screening which eliminated participants who "(a) failed to complete or inaccurately completed the training or assessment tasks (n = 34); (b) failed the prelearning vocabulary test (n = 8); (c) indicated the animacy rule, either online or offline (n = 9); or (d) clearly demonstrated a non-animacy-based strategy that would result in obtaining 50% accuracy on the critical items on the assessment tasks (n = 11)." (p. 473). The authors allow that "[i]t may be argued that eliminating these participants from the study for noncompliance raises the level of internal validity in the study." (p. 488).

[5] Rule awareness was assessed through a post-experiment structured interview wherein participants were first asked what criteria they had used to make their decisions during the testing phase, and then, more specifically, how or when they used each of the novel words. If animacy was mentioned, the participant was asked at what point in the

forced-choice task) was found in either of these groups, going against the findings of rule-unaware learning reported by Williams (2004, 2005). Previous knowledge of gendered languages was not found to be a significant factor in performance (in contrast to the results of Williams, 2004; 2005), but the authors note that their reported lack of variation between participants in terms of familiarity with gendered languages would make it unlikely to pull out statistically significant effects based on individual factors.[6] Faretta-Stutenberg and Morgan-Short write that "it is difficult to pinpoint the exact cause for inconsistent results between the present study and Williams (2005)" (2011, p. 27). However, they note differences between their participant sample and that of Williams (2005) in terms of overall linguistic knowledge, fields of study, years of education, and age, which might explain the disparity in results regarding the possibility of L2 morphosyntax learning without awareness (2011, p. 27).

In another study that aimed to investigate L2 morphosyntax learning without awareness, Leung and Williams (2011) used reaction time measures rather than a forced-choice decision task as a potentially more sensitive measure of implicit learning. This sought to address Hama and Leow's (2010) reported failure to replicate the findings of implicit L2 learning reported in Williams (2004, 2005). The authors' reaction-time approach was inspired by methodologies in cognitive psychology for studying topics such as contextual cuing (Chun, 2000; Jiang & Chun,

study they became aware of the relevance of animacy. Otherwise, the participant was asked whether they had looked for rules regarding the novel words at any point in the study. Finally, the participant was told that there was a rule used to govern the use of the determiners, and asked to guess as to what the rule was. Participants who did not report awareness information at this point then performed a second testing phase, but this time with the goal of trying to figure out the rule that dictated the appropriate word to end each sentence. This second test phase was followed by a second set of debriefing questions in which participants were then asked to state the underlying rule that they had noticed. Participants who did not report the correct rule were then asked if they had considered over the course of the experiment that some words were used with living things and others with inanimate objects.

[6] Specifically, the authors write that "due to a lower mean number of gendered languages known among participants in the present study (0.80) in comparison with participants in Williams (2005) (1.68), it is still not possible to determine whether the study of language related disciplines or knowledge of gendered languages is more relevant to implicit learning of the determiner system under investigation" (Faretta-Stutenberg & Morgan-Short, 2011, p. 27).

2003) and derived attention (Lambert, 2002; Lambert & Sumich, 1996), in which participants' fine-grained reaction time differences to trials that do vs. do not follow an underlying pattern are taken as a sensitive operationalization of implicit learning. This reaction time difference has been attributed (e.g., by Batterink et al., 2014) to interference effects not unlike those involved in the famous Stroop paradigm (e.g., MacLeod, 1991). Leung and Williams (2012) argued that introducing time pressure in the context of a reaction time task would provide a more sensitive measure for the study of L2 learning without awareness because it would encourage implicit rather than explicit processes, as found by both non-linguistic (Destrebecqz & Cleeremans, 2001) and linguistic (R. Ellis, 2005) research. An additional advantage of such a reaction time-based methodology over the previously used forced choice decision tasks is that, because the measure is concurrent over the period of learning, then from the participants' point of view there would be no division between the experiment's training and test phases. This would make it easier for implicit learning to be detected because an abrupt change of the task requirements (as in the transition from a sentence-reading task to a forced-choice task in the preceding studies) would work against the inflexible and context-specific nature of implicit learning (Roediger, 1990; Jiménez et al., 2006). Another advantage of the reaction-time methodology is that the experimental task would involve comprehension rather than production, meaning that lower processing demands would be imposed on the learner, making it easier for implicit knowledge to emerge (Hulstijn, 2002). Note also that, by avoiding a forced choice task, the authors circumvented potential critiques of artificial language studies that argue that chance levels are problematic to define (e.g., Hamrick & Sachs, 2018).

In Leung and Williams's (2011) study introducing this reaction time method, participants were 25 English native speakers with a variety of L2s, although "[n]o participant had advanced knowledge of any languages with highly developed case systems, such as German or Latin" (p. 45). As with the previous experiments, participants were told only one dimension of the underlying rules governing the use of the artificial language articles *ul*, *gi*, *ro*, and *ne*. This experiment used the age of a referent (i.e., whether the referent was an adult or child) as the overt dimension and whether the referent was a semantic agent or patient (broadly speaking, the doer or recipient of an action) as the covert dimension. Participants performed an oral picture description task; a reaction time task wherein participants indicated whether a referent appeared on the left or right side of a picture (which could be facilitated by knowledge of the covert agent/patient rule); and a sentence reformulation task wherein participants produced the sentence in rule-adhering English word order. When analyzing the reaction time responses of the 20 participants who reported no awareness during a post-experiment interview, the authors found that response times were significantly slower to trials that violated the hidden animacy regularity (presented in a "violation block" at the end of the experimental phase). As such, this reaction time-based methodology corroborated findings of morphosyntactic learning without awareness from the forced choice tasks used in Williams (2004, 2005).

Leung and Williams (2012, Experiment 1) performed a subsequent study that also examined rule-unaware morphosyntax learning using reaction times but with a slightly simplified paradigm. In this study, participants were 33 native English-speaking students at the University of Cambridge with a variety of prior second language experience. The stimuli consisted solely of the artificial language articles *ul*, *gi*, *ro*, and *ne* paired with an accompanying noun. Participants were told that two of the articles denoted nearby referents and that the other two denoted distant referents. However, they were not told that an underlying animacy rule was involved. In the main experimental phase, participants would see a picture with two possible referents; hear an accompanying audio description of one of the referents (e.g., "ro bull"); decide whether the named referent was living or nonliving; and then orally repeat and then translate what they heard (e.g., "ro bull, the far bull"). Awareness of the hidden animacy rule was assessed via a post-experiment interview that gradually asked more and more specific questions regarding the underlying animacy rule.[7] The twenty participants who demonstrated no rule awareness in this interview showed slower reaction times as well as a higher error rate in a violation block towards the end of the experiment in which the artificial language articles' animacy assignments were reversed, suggesting that L2 morphosyntax acquisition without awareness is possible.

Subsequent studies have used this reaction-time measure coupled with post-test questionnaires to test implicit learning for domains beyond animacy and subject/object thematic roles. For instance, one study found that learners could also implicitly learn the semantic preferences for novel verbs, i.e., whether they took on abstract or concrete nouns as their objects (Paciorek & Williams, 2015). However, findings of implicit learning did not hold in subsequent experiments for the encoding of a size-based semantic distinction (Leung & Williams, 2012, Experiment 2) or for supposedly "linguistically anomalous" concepts like the number of capital letters in an English word or the number of strokes in a Chinese character (Leung & Williams, 2014, p. 1). These results suggest that implicit learning is limited to some extent by the nature of the meanings involved. Furthermore, there is also evidence that such implicit learning could

[7] The interview procedure used in Leung and Williams (2012) is described by the authors as follows (p. 13): "Participants were firstly asked what they thought the experiment was about. Then participants were asked whether they noticed anything odd toward the end of the experiment and if so, what the oddity was. If the participants were aware of the association between the determiners and the animacy system and its violation, they were likely to report their observation sat this point. Participants were then asked whether they had any feelings about the different conditions in which gi versus ul (both meant near) and ro versus ne (both meant far) were used. They were asked to make as many guesses as they could and state whether they were confident in their guesses. They were allowed to look at the training materials in order to facilitate their guess. Participants were classified as “unaware” if they did not show any knowledge of the correlation between the use of determiners and animacy and were unable to match the determiners with their animacy values."

differ across languages: in another study using the same reaction time methodology and awareness assessment (Leung & Williams, 2014), both Cantonese native speakers and English native speakers could implicitly learn an animacy-based distinction (replicating Leung & Williams, 2011), but only Cantonese speakers could implicitly learn a shape-based distinction related to a Cantonese classifier system. Such results lend support to original findings of L2 morphosyntax learning without awareness while providing insight into the particular nature of the possible encoded meanings that are and are not susceptible to such learning.

In a more recent extension study that sought to address critiques of think-aloud and post-test measures of rule awareness (further discussed below in Section 2.2.4), Rebuschat and colleagues (2013) recreated Williams' (2005) artificial language experiment but sought to avoid relying on verbalization to distinguish implicit from explicit processes. Instead, the authors measured participants' awareness by collecting per-trial measures of participants' subjective confidence in response accuracy as well as the perceived source of their response, from a choice of *guess*, *intuition*, *memory*, or *rule knowledge*. At the end of the experiment, a post-test debriefing questionnaire was also performed to specifically assess awareness of the underlying animacy rule (as in Faretta-Stutenberg & Morgan-Short, 2011; Hama & Leow, 2010). Participants for this study were 30 native English-speaking undergraduate students with a variety of second languages; 15 participants were assigned to an experimental group and 15 to a trained control group.[8] The authors replicated Williams (2005) in finding signs of learning (i.e., above-chance accuracy on a two-alternative forced choice task) for trials that had received source

[8] The trained control groups in Rebuschat et al. (2013) received the same training conditions as the experimental groups, but with the mapping between the novel words and the relevant animacy feature randomized and balanced. This trained control condition was included so as to address the possibility that a control group would have performed better than chance levels simply due to response biases based on prior knowledge (e.g., Reber & Perruchet, 2003).

attribution responses of either *guess* or *intuition*, which would entail implicit rather than explicit processing (Rebuschat et al., 2013).

2.2.4 Issues with a behavioral-only paradigm

There are several potential issues with the assessments of awareness in the studies discussed so far. For instance, Rebuschat and colleagues (2013) argue that self-report measures of awareness (whether collected through concurrent think-aloud verbalizations or through post-experiment questionnaires) rely on two unwarranted assumptions: that all learning is unaware if participants do not verbalize relevant features, and that knowledge is unconscious when participants show gains from training despite being able to verbalize the relevant knowledge. They note that participants' memory of their conscious thoughts may degrade by the time of a post-testing questionnaire, and that verbalization may not capture all that one becomes aware of given that "subjective awareness is fleeting and cannot be completely recorded" (Schmidt, 1995, p. 28, quoted in Rebuschat et al., 2013, p. 259). Furthermore, Rebuschat and colleagues (2013) write that participants may not verbalize knowledge that they are unsure of or do not realize is relevant. This is corroborated by findings that participants often report things on a second attempt at verbalization that they do not report in the first (Erdelyi & Becker, 1974, cited in Dienes & Berry, 1997).

Beyond the possibility of under-reporting (e.g., if participants forget or hold back a full oral report response because of limited confidence), another possibility is that participants over-report their level of knowledge, e.g., if they deduce the rules during a post-test questionnaire, such that their oral report does not correspond to the processes actually used during the prior experiment (Hama & Leow, 2010). Attempting to avoid these issues by employing concurrent verbal reports (i.e., "think-aloud" protocols) is also problematic because of the potential of reactivity, i.e., the possibility that participants' task performance may be affected by the reporting task (Bowles, 2010). For instance, thinking aloud could affect participants' L2 cognitive processing due to the added processing load from the secondary task (Leow & Morgan-Short, 2004). This has been demonstrated in one study which found that only a silent, non-think-aloud participant group was able to generalize knowledge to new test items, in contrast to groups who thought aloud during training or during both training and testing (Rebuschat et al., 2015). In this way, verbal reports (either during or after an experimental task) may provide only limited measures of rule awareness.

Although Rebuschat et al. (2013) used source attribution ratings to avoid these limitations, this methodology is not unproblematic, either. One issue with the use of subjective measures is that they have been found to induce participants to rely on grammar rules rather than knowledge of chunks in the context of an artificial grammar task (Invachei & Moroshkina, 2018), perhaps by suggesting to participants different strategies that could be taken or by making participants unduly self-conscious of the strategies that they take. In this way, subjective measures-based approaches may introduce the very sorts of reactivity issues that they were designed to avoid. Although this problem can be mitigated by "triangulating" learner knowledge from a variety of measures (Godfroid & Schmidtke, 2013; Rebuschat et al., 2015), this would not provide a single objective measure of awareness, in that data would ultimately be tainted to some extent by the methodologies used to assess rule-awareness and explicit processing. It would seem, then, that behavioral measures alone are limited for studying the roles of awareness in L2 morphosyntax learning.

By way of interim summary, previous SLA studies using an artificial language paradigm introduced by Williams (2004, 2005) paradigm have found evidence for learning in the absence of rule awareness (Leung & Williams, 2011, 2012; Rebuschat et al., 2013). However, issues with the behavioral methodologies limit researchers' confidence in the operationalizations of implicit and explicit learning, which may explain failures to replicate findings of learning without awareness (Hama & Leow, 2010; Faretta-Stutenberg & Morgan-Short, 2011). These issues are captured by Leow below:

> At the non-concurrent reconstruction stage, limitations include the inability to (1) methodologically establish participants' behavior during the experimental phase of the study…, (2) ascertain whether the offline performance reflects accurately the learning behavior of each experimental learning condition (did some participants in the implicit or incidental learning condition become aware of the target L2 information?), and, more specifically, (3) gather data on how participants actually processed (low or higher depth of processing) the target information (2018, p. 7).

Almost three decades ago, Schmidt wrote that:

> There is almost a complete lack of evidence in the second language literature which is directly relevant to the [noticing] hypothesis, since second language researchers have never asked learners to provide systematic information on what they notice while learning languages that could be compared to what they can be shown (by other measures) to have learned. (1990, p. 139-140).

If I replaced "asked learners to provide systematic information on what they notice" with "taken direct and objective measures of what learners notice," then this statement would seem to hold true today.

2.2.5 Using EEG to overcome limitations with behavioral measures

One way to address the issues described above that arise when using behavioral measures to investigate implicit vs. explicit L2 morphosyntax learning is through electroencephalography (EEG). As reviewed by Morgan-Short (2014), EEG presents several potential benefits over behavioral methodologies because it can detect neural activity associated with language processing in a covert and non-invasive way at a fine level of temporal resolution. More specifically for the context of this dissertation, one could examine neural firing that characterizes language processing with different levels of awareness by measuring event-related potentials (ERPs), or the observed deflections in amplitude evoked by different experimental conditions. Different ERP components have been identified for stimulus processing involving different levels of consciousness (for a review, see Rutiuku & Bachman, 2017), which may align with previous SLA models' distinctions between levels of attention, e.g., Leow's (2015) descriptions of attended, detected, and noticed input. For instance, Batterink and Neville (2013) found a difference in ERPs between consciously perceived L1 grammatical violations (which elicited a P600 component) and unconsciously perceived violations (which elicited a left anterior negativity). Such findings suggest that EEG can be used to detect differences in L2 morphosyntax processing between (purportedly) rule-aware vs. rule-unaware learners, which would support the validity of the measures of awareness taken by the experimenters.

Precisely such a neurolinguistic study was performed by Batterink et al. (2014), who collected EEG data in an experiment that adopted the reaction-time artificial language paradigm used by Leung and Williams (2012). This involved exposing participants to word couplets consisting of nouns preceded by a novel artificial language article (either *ul*, *gi*, *ro*, or *ne*) which followed an overt near/far distance rule as well as a covert animacy rule. After a brief pre-training phase to introduce these novel pseudowords, participants (29 native English speakers) performed a task wherein they saw a pseudoword-noun couplet and were tasked with judging first whether the referent was living or nonliving and secondly whether the referent was near or far; this ensured that the novel pseudoword and the notion of animacy were processed. Rule awareness was assessed in a structured interview following the main experimental phase as follows: participants were first asked if they had noticed any pattern about when the different articles were used, beyond the overtly taught near/far rule. If at this point participants spontaneously reported that certain articles co-occurred with living/nonliving referents more often than others, they were asked at what point in the experiment they had noticed this pattern. If participants did not report the relevant pattern, they were asked whether they looked for rules/patterns during the experiment, and then to guess which underlying rule determined article usage. If at this point participants still did not list animacy as the correct rule, they were then invited to guess which articles co-occurred more often with living vs. non-living nouns. Participants who produced the correct pattern and reported having noticed it during either the first or second experimental block were classified as rule-aware. Otherwise, they were classified as rule-unaware.

Besides the simultaneous collection of EEG data, Batterink et al.'s (2014) study featured several other important differences with the methodology used by Leung and Williams (2011). Firstly, rather than including a violation block at the end of experiment, in Batterink et al.’s (2014) experiment one of every seven trials contained a violation to the pattern, such that reaction time slowdowns to the animacy response (indicating learning of the hidden regularity) could be assessed over the course of the experiment rather than only at the end. Secondly, a 90-minute nap phase was introduced in the middle of the experiment, such that the relative effects of sleep on rule-aware vs. rule-unaware participants’ learning could be determined.

In Batterink et al. (2014)’s results, reaction time slowdowns to trials that violated the underlying animacy rule indicated learning in both rule-aware and rule-unaware participants, replicating previous findings of implicit morphosyntax acquisition (e.g., Leung & Williams, 2012). This was corroborated by an analysis of accuracy differences for rule-adhering vs. rule-violating trials. Correlations from sleep data (which was collected between two training phases) indirectly suggested that a similar underlying neural mechanism was involved between rule-aware and rule-unaware participants, because sleep had similar effects of improving learning for both groups. However, ERP differences between the rule-aware and rule-unaware participants indicated that the two groups were actually processing the stimuli differently: in response to violation trials, rule-unaware participants showed a negative deflection in the EEG signal in an early time window (400-800 milliseconds following the presentation of the article-noun couplet), whereas rule-aware participants showed a positive deflection in a later time window (800-1100 ms post-stimulus). These findings suggest that the verbal reports from the post-experiment questionnaire were indeed capturing a difference between how these two participant groups processed the linguistic task. This addresses (at least in terms of group-level effects) the potential critique that participants who were classified as rule-unaware were simply underreporting their level of explicit knowledge, or that participants classified as rule-aware were in fact rule-unaware during the bulk of the experiment itself (e.g., Hama & Leow, 2010; Rebuschat et al., 2015). In this way, findings from the EEG experiment reported in Batterink et al. (2014) increase my confidence in previously reported findings of L2 morphosyntax learning without rule awareness (e.g., Leung & Williams, 2011).

As indirect support for their findings, Batterink et al.'s (2014) observed ERP results—namely, a negativity to implicit rule violation processing and a positive component to conscious rule violation processing—parallels neatly with other EEG studies that touch indirectly on the role of implicit vs. explicit processes in L2 learning. In the first of these, Morgan-Short and colleagues (2012) found different ERP signatures for artificial language learner groups that were trained in either implicit vs. explicit instruction conditions. Although the two groups were similar in terms of their behavioral performance, participants who received sheer exposure to the language (an implicit condition) showed an N400 (an early negative component previously associated with violations of expectation; Kutas & Federmeier, 2011) in response to grammatical violations at low proficiency and a (native-like) pattern of an anterior negativity followed by a P600 ERP component and a late anterior negativity at high proficiency. By contrast, participants who received explicit instruction on the artificial language's grammar rules showed no significant effects at low proficiency and only an anterior positivity followed by a P600 at high proficiency. Although this study manipulated training conditions rather than collecting by-participant or by-trial measures of implicitness/explicitness (meaning that the learners' exact levels of rule awareness could not be inferred), Morgan-Short and colleagues' (2012) finding of ERP differences in L2 morphosyntax processing across more vs. less explicit conditions parallels the differences found by Batterink et al. (2014) across rule-aware and rule-unaware learners.

In a second study that indirectly lends support to Batterink et al.'s (2014) findings, Wan et al. (2010) used participants' subjective reports of response source attribution to find different EEG responses to artificial grammar violation trials when these involved more explicit processing (i.e., when the participant's response was attributed to rule knowledge or to recollection of a specific memory) vs. more implicit processing (i.e., when the response was attributed to a guess, intuition, or general familiarity with the test item). In the overall results, comparisons of all grammatical vs. ungrammatical trials regardless of source attribution revealed an N2 ERP component, which is a negative deflection that is indicative of surprisal due to unexpected stimuli and thus an index of learning. Meanwhile, the P300 component (indicative of conscious processing) was larger for trials that were attributed to conscious sources (i.e., rules or memories) as opposed to unconscious ones (i.e., guesses, intuition, and familiarity). As such, Wan et al.'s (2010) findings coincide with Morgan-Short et al.'s (2012) results to indicate differences in neural activity during more vs. less explicit L2 rule-violation processing, just as in the Batterink et al. (2014) study.

2.2.6 Motivation for replication of Batterink et al. (2014)

To this author's knowledge, Batterink et al. (2014) represents the first and only experiment to use EEG methods to compare L2 morphosyntax learning in rule-aware vs. rule-unaware participants. Batterink and colleagues' study addresses critiques of previous behavioral methodologies by reporting evidence of implicit L2 learning at the neural level, thus lending support to the measures used to establish rule awareness. These findings are highly relevant for psycholinguistic debates between competing models of attention and awareness in SLA (e.g., Schmidt, 1990; Tomlin & Villa, 1994; Robinson, 1995; Leow, 2015). In addition, this study may have implications for L2 pedagogy in that it shows a qualitative difference between learning that involves subliminal learning from sheer exposure (as in naturalistic learning conditions) as opposed to willful learning that involves conscious awareness (as in metalinguistic teaching approaches).

Given its innovative multidisciplinary use of EEG as well as its relevance for psycholinguistic theory and L2 pedagogy, Batterink et al. (2014) is an important study to replicate. This is particularly true in light of the growing movement for more replication studies in the fields of psychology (Lamal, 1990; Francis, 2012) as well as second language acquisition (Marsden et al., 2018). A replication would also be especially important given the call for increased standards for EEG research (Keil et al., 2014) and the substantial possibility of finding spurious significant results when using EEG methodologies (Luck & Gaspelin, 2017). As such, a replication of Batterink et al. (2014) would contribute to my understanding of the role of awareness in L2 acquisition by increasing my confidence in its reported findings. Furthermore, for the purposes of this dissertation, a replication experiment would form the basis of several extension analyses as described in Section 2.3.5.

Although the main research question in Batterink et al. (2014) was framed around the role of sleep in morphosyntax acquisition, for the purposes of this dissertation it is reworded more broadly for a second language acquisition context as follows:

Research Question 1: Can event-related potentials capture differences in learning of a covert morphosyntactic regularity between rule-aware and rule-unaware learners?

2.3 The interplay of implicit vs. explicit knowledge and processes

Going beyond the question of whether implicit learning of an L2 morphosyntactic rule is possible at all, this section focuses on the nature of the interplay between implicit and explicit learning within individual learners. This follows Simard and Wong's suggestion that, "instead of asking whether attention and awareness are necessary or not for SLA, it may be more fruitful for research to examine how different levels of attention and awareness impact learning" (2001, p. 120-121). To illustrate this, it is unclear whether explicit learning interferes with implicit learning; has no relationship with explicit learning; or facilitates implicit learning. More specifically, the questions discussed in this section include: do operationalizations of learning reflect the same underlying process for rule-aware vs. rule-unaware participants? Do implicit and explicit routes to acquisition overlap in the kinds of neural mechanisms that they induce, or do they involve entirely distinct neural mechanisms? Does having awareness of a rule help or hinder learners from acquiring implicit-like brain responses? Can only one process or one sort of knowledge be in play at any given time, or can they co-occur? If so, do these two forms of learning interact with each other in any way? The answers to such questions would be highly informative for L2 teaching praxis, in that they would lend insight into the relative costs and benefits of one processing style (and by extension, one instructional style) over another.

2.3.1 Theories on the interface between implicit and explicit processing in SLA

Previous models in the field of second language acquisition have argued alternatively for a strong interface, a weak interface, or no interface between implicit and explicit L2 learning. These frameworks do not seek to describe whether explicit knowledge is necessary for acquisition in the first place, but rather are oriented around what the effects of explicit learning on implicit learning might be. Each of these positions is described below.

One theoretical position has argued for a strong interface between implicit and explicit processes (DeKeyser, 1997, 1998, 2003; O'Malley, Chamot & Walker, 1987; Sharwood-Smith, 1988). The most prominent model in this camp comes from DeKeyser's Skill Acquisition theory (2007), which traces its origins to the field of cognitive psychology (e.g., Anderson, 1982). According to this theory, language learning (and any cognitive skill in general) consists of and proceeds through a series of stages. At an initial declarative stage, learners accumulate a factual understanding of L2 forms ("knowledge that"), such as verbalizable explanations of grammatical rules. This explicit knowledge may be acquired through overt instruction (e.g., explicit rule presentation and focused practice), but it might also result from simple exposure or observation on behalf of the learner. At a subsequent procedural stage, learners act on this declarative knowledge ("knowledge how") to actually process and produce language. After such gradual proceduralization (over the course of extended communicative practice), learners finally arrive at a stage of automatization, wherein procedural knowledge gradually becomes effortless, spontaneous, and fluent. At this final stage, explicit knowledge becomes "functionally equivalent to implicitly acquired knowledge" (DeKeyser & Juffs, 2005, p. 461). The strong interface position overtly states that explicit knowledge "is conducive or plays a causal role" in the acquisition of "procedural, automatized, or implicit knowledge" (DeKeyser, 2009, p. 126).

In contrast to DeKeyser's (2009) strong interface position is the weak interface position, according to which explicit knowledge does not always lead to the acquisition of implicit knowledge. Two different weak interface positions have been described, entailing either an incomplete connection between implicit and explicit knowledge systems or an indirect relationship between explicit processing and learning. The former position is taken by R. Ellis (1994, 2004, 2005), who argues that both implicit and explicit knowledge are possible outcomes of instructed SLA but that the facilitative effect of explicit knowledge for the learner's implicit systems can only occur for certain aspects of the L2 system. This depends on the nature of the grammatical elements involved and particularly on whether these linguistic forms are developmentally constrained. For so-called developmental elements, explicit knowledge becomes implicit only when learners are developmentally ready. Conversely, explicit knowledge of non-developmental items can become implicit at any time. In this way, under R. Ellis's approach, the link between explicit and implicit knowledge is incomplete in that explicit knowledge cannot always become implicit knowledge, based on the nature of the forms involved. This contrasts with the causal role of explicit knowledge put forth in DeKeyser's (2007) more deterministic strong interface approach.

In another variation of the weak interface hypothesis, N. Ellis (2005, 2006, 2007) argues that L2 learning is at heart an inherently implicit process that is driven by gradual associative acquisition of form-function mappings, based on probabilistic encounters with relevant exemplars. Although learning is itself implicit, explicit knowledge and processing can influence acquisition through indirect means. For instance, explicit instruction and feedback can guide learners to relevant forms so as to maximize implicit learning. Explicit instruction can also help to overcome interference from learned attention, i.e., attentional routines employed during language processing that are carried over from the learner's first language and that may interfere with the (underlyingly implicit) processing of L2 input. This view of explicit consciousness as potentially affecting but not directly driving L2 acquisition aligns with the so-called "associative-cognitive creed" in SLA (VanPatten & Williams, 2007) and stems more generally from usage-based views on L2 development (e.g., Goldberg, 2006; Bybee, 2008; Langacker, 2009). What distinguishes this version of the weak interface position from the strong interface position is its claim that explicit knowledge is helpful for implicit learning but does not play a strictly causal role (Hulstijn, 2015, p. 36)

In contrast to the above approaches, the no-interface hypothesis posits that implicit and explicit L2 knowledge are inherently different in that they are acquired through different means (Hulstijn, 2002; Krashen, 1981), are located in and/or accessed by different parts of the brain (Paradis, 1994), and are invoked by different processes (R. Ellis, 1993). Given this inherent distinction, explicit knowledge is a separate type of knowledge and can thus never become implicit. As such, explicit knowledge will always remain explicit regardless of the amount of L2 exposure, practice, or proficiency. In this way, the no-interface hypothesis denies that explicit knowledge can facilitate the acquisition of implicit knowledge (Krashen, 1981, 1985). Another prediction of the no-interface position is that not every L2 linguistic form is learnable (Krashen, 1981; Schwartz, 1993; Truscott, 1996; Paradis, 2009). This hypothesis argues that conscious learning cannot make up for incomplete acquisition; that language is too complex to be explained/learned explicitly; and that what is learned explicitly cannot be deployed by learners in authentic, spontaneous communication. As such, explicit instruction may only help in terms of externally observable performance when learners are in closely monitored conditions, but not in terms of helping to develop the sort of underlying linguistic competence that drives naturalistic and spontaneous L2 processing and production in more authentic settings. In this way, this no-interface position captures the possibility that explicit learning can occur without any consequences for implicit learning.

In sum, three prominent viewpoints regarding the interface of implicit and explicit processing and knowledge in L2 acquisition are that explicit knowledge always helps (strong interface hypothesis), sometimes helps (weak interface hypothesis), or does not help (no-interface hypothesis) the acquisition of implicit knowledge. Another possibility that has not previously been presented formally as a theoretical position in the interface debate is that explicit knowledge can hinder the development of implicit knowledge. Such a situation might arise in the case of *blocking*, wherein consciously allocated attention to certain learned cues in the L2 input distracts learners from acquiring other cues. Indeed, previous studies in SLA have reported that L2 forms that are acquired at an earlier stage can effectively "block" learning of subsequent forms due to competition for attention (N. Ellis, 2006; Ellis & Sagarra, 2010; Solman & Chung, 1996). As a more specific illustration of how explicit processing can hinder learning, Leow notes that overtly paid attention can have inhibitory consequences, e.g., when attention is aligned to an incorrect form or when perceived linguistic input has a strong mismatch with expectations (1998, p. 146). The following section presents empirical research in the field of SLA that is relevant to these different positions on the interplay of implicit and explicit L2 learning.

2.3.2 Empirical studies in SLA regarding the implicit/explicit interface

What do previous experiments in L2 acquisition say about the competing interface hypotheses? Although research that specifically addresses this topic is relatively scarce (as noted by Suzuki & DeKeyser, 2017), patterns in SLA findings to date indicate that instruction is helpful "to some extent, for some forms, for some students, at some point in the learning process" (DeKeyser, 1998, p. 42), as borne out by the meta-analytic research described in Section 2.2.2 (Norres & Ortega, 2000; Spada & Tomita, 2010; Goo et al., 2015). Such findings suggesting that explicit processing and knowledge make at least some positive contribution to linguistic gains, even in spontaneous language use (R. Ellis, 2002; Russel & Spada, 2006), would suggest that there is at least *some* interface between the implicit and explicit domains, providing potential evidence against the no-interface hypothesis.

As has been noted before, one major challenge to drawing conclusions about the interplay of implicit vs. explicit knowledge and processing from the SLA research included in these meta-analyses—and particularly for early studies on this topic (e.g., Hulstijn & Hulstijn, 1984; Seliger, 1979; Sorace, 1985)—is the possibility that implicit conditions actually generated explicit knowledge. This has been reported in several experiments wherein there were no specific instructions to learn any particular information and no information regarding a post-exposure test (Grey, Williams, & Rebuschat, 2014; Hamrick & Rebuschat, 2012, 2014; Rebuschat & Williams, 2009, 2012; Tagarelli et al., 2011; Hamrick, 2014; Rebuschat et al., 2013, 2015; Rogers et al., 2016). Compounding this issue, a higher amount of implicit exposure has been tied to a higher chance of acquiring explicit knowledge: in an artificial grammar learning study (a methodology further described in Section 2.3.4), Mathews et al. (1989) found that participants who first performed an implicit training task were more likely to become aware of the rule system in a subsequent rule discovery task than participants with no prior exposure to the target stimuli. It is not hard to imagine why explicit processes may come from implicit-like conditions: given what Wonnacott et al. call the "experimental pragmatics" (2012, p. 475) of laboratory-based SLA studies, participants do not enter experimental conditions without at least the minimal intention of learning something. As Reber notes (1993, p. 26), “if participants feel they can ‘crack the code,’ they will attempt to do so.” An illustration of why this blurring of the implicit/explicit line is a problem for interface debates comes from Batterink and Neville (2013), who found that both implicitly- and explicitly-trained learners showed (native-like) P600 ERP responses to grammatical violations on an artificial language syntactic rule, with the magnitude of this effect tied to participants’ behavioral proficiency with this rule. However, post-test questionnaires showed that successful learners from both the implicitly and explicitly trained groups were able to explicitly verbalize this L2 rule. This suggests that differences between training conditions were not necessarily borne out in participants’ internal cognitive behavior. In other words, learners who acquire explicit knowledge from an implicit training condition may as well have been in an explicit condition. This limits the generalizability of findings from studies that manipulate the implicitness/explicitness of training *conditions* for making claims about interfaces between implicit and explicit knowledge and processes.

The difficulties in measuring implicit knowledge and processing as described above have motivated the use of carefully controlled methodologies with more sensitive measures of learning that utilize online (i.e., real-time) measures of comprehension. For instance, the use of behavioral response times and eye-tracking in the context of language comprehension tasks can be used to assess implicit knowledge because they prevent L2 learners from processing explicitly (Suzuki & DeKeyser, 2015; Suzuki, 2017). Such methods leave “virtually no room” (Suzuki & DeKeyser, 2017, p. 752) for explicit knowledge to be accessed in a conscious manner because they can be measured the scale of hundreds of milliseconds (Andringa & Curcic, 2015; Dussias et al., 2013; Godfroid, 2016; Suzuki, 2017; Suzuki & DeKeyser, 2015; Vafaee et al., 2017). By contrast, form-focused tasks such as timed grammaticality judgments arguably involve more conscious processes because these require participants to reflect on metalinguistic knowledge (DeKeyser, 2003; Vafaee et al., 2017). As such, online psycholinguistic methodologies can be leveraged to assess implicitness and explicitness in a more controlled fashion than studies that merely manipulate learning conditions.

In one eye-tracking study by Cintrón-Valentín and Ellis that explored the interface of implicit and explicit processing (2015), English and Chinese native speakers read Latin sentences wherein temporality (i.e., past vs. present tense) could be determined either through (adverb-based) lexical cues or (conjugation-based) morphological cues. Participants were divided into different training conditions, with either explicit metalinguistic explanations; bolding and highlighting (but no explanation) of relevant forms; verb translation pre-training; or a control condition. The researchers found that the relative amount of eye fixations to linguistic forms during input processing was correlated with successful use of these cues in real-time comprehension and production. Furthermore, explicit instruction was found to lead participants to better use of such cues (both in comprehension and production), allowing learners to overcome previous biases from their L1 (i.e., against the use of morphological cues in Chinese native speakers). These results suggest that explicit form-focused instruction may lead to “hardwiring” of associations and/or input processing routines that would otherwise only come as the outcome of gradual implicit acquisition, described by the authors as a slow and purely statistical process. More broadly for the purposes of the dissertation, these results suggest that explicit processes of attention allocation can facilitate implicit learning, thus supporting the weak and strong interface hypotheses.

Some potential evidence against the strong interface hypothesis comes from Andringa and Curcic (2015), who aimed to “find evidence of either conversion or application of explicit knowledge in an online processing task favoring the use of implicit knowledge” (p. 238). The authors performed an eye-tracking study to determine whether explicit instruction could help learners to circumvent implicit statistical learning, operationalized here as eye movements indicative of prediction of upcoming linguistic input in a visual world (i.e., scene-viewing) paradigm. This would test the strong interface hypothesis’s claim that explicit learning directly affects implicit learning. The authors found that metalinguistic instruction on a differential object marking rule (involving differences in preposition use between animate and inanimate objects) led to improved performance in an online grammaticality judgment task, relative to a control group that received only exposure and no metalinguistic instruction. However, the eye-tracking data suggested that explicit instruction did not lead to differences in eye movement behavior (interpreted by the authors as a measure of implicit processing). This would seem to present evidence against predictions of a positive effect of explicit processing on L2 acquisition as posited by the strong interface hypothesis and (to a lesser extent) by the weak interface hypothesis.

Another study that speaks to the interface of implicit and explicit learning in second language acquisition comes from Suzuki and DeKeyser (2017), who assessed a cohort of advanced Japanese L2 speakers in terms of implicit knowledge (assessed via a visual-world task, a word monitoring task, and a self-paced reading task), explicit knowledge (assessed via a timed auditory grammaticality judgment task, a timed visual grammaticality judgment task, and a timed fill-in-the-blank task), explicit learning aptitude (assessed via an adapted version of the LLAMA F test, which involves inferring grammatical rules based on picture and word sequences; Meara, 2005), and implicit learning aptitude (assessed via a serial reaction time task, a method described more fully in Section 2.3.3). Using structural equation modeling, the researchers found that explicit learning aptitude significantly predicted acquisition of explicit knowledge, which in turn significantly predicted acquisition of implicit knowledge. As such, Suzuki and DeKeyser's (2017) results support the notion of a facilitative role for explicit knowledge on implicit knowledge, thus providing evidence in favor of the strong and weak interface hypotheses.

In sum, the three studies described above (Andringa & Curcic, 2015; Andringa & Curcic, 2015; Suzuki & DeKeyser, 2017) go beyond the manipulation of the implicitness/explicitness of training conditions in their use of more controlled experimental measures for assessing the interplay of implicit and explicit learning. They align with meta-analytic results (e.g., Norris & Ortega, 2000; Spada & Tomita, 2010; Goo et al., 2015) in suggesting that explicit knowledge can facilitate the acquisition of implicit knowledge. However, the findings of learning in the absence of rule awareness reviewed previously in Section 2.2.3 (e.g., Williams, 2004, 2005; Leung & Williams, 2011, 2012, 2014; Rebuschat et al., 2013) suggest that explicit knowledge may not be strictly necessary for acquisition to occur. In order to provide a broader perspective regarding the relationship between implicit and explicit learning, the following section reviews studies from the wider field of cognitive psychology that have investigated this topic through non-linguistic methodologies.

2.3.3 Findings from psychology experiments on implicit vs. explicit processing

What do previous studies from the field of cognitive psychology tell us about the interface (if any) between implicit and explicit processes? A long line of prior research has found a role for implicit knowledge (see Cleeremans et al., 1998; Perruchet, 2008; Reber, 1989; Shanks, 2005, for reviews) for learning in many disparate domains, including social interaction

(Lewicki, 1986), music perception (Dienes & Longuet-Higgins, 2004; Rohrmeier et al., 2011), intuitive decision making (Plessner et al., 2008), and—most importantly for the purposes of this dissertation—language comprehension and production (Berry & Dienes, 1993; Reber, 1993; Williams, 2009). Conversely, an important role has also been found in learning for explicit factors like attention, awareness, strategies, and declarative memory (Redding & Wallace, 1996; Hwang et al., 2006; Michel et al., 2007; Taylor & Thoroughman, 2007, 2008). Beyond just formal academic studies, the distinction between implicit and explicit components in the field of psychology has also been reflected in publications aimed for broader audiences (e.g., the popular non-fiction work *Thinking, fast and slow*; Kahneman & Egan, 2011). These typically propose dual-process and/or dual-system views, wherein automatic processes generate impressions and tentative judgments that may be blocked, accepted, or corrected by more controlled processes (Morewedge & Kahneman, 2010). In more recent years, neuroimaging approaches have allowed researchers to explore the neural correlates for each of these different kinds of processes (Smith et al., 2006; Galea et al., 2011; Choi et al., 2014; Yang & Lisberger, 2014). Although some principal findings are discussed below, a more comprehensive review of non-linguistic studies on implicit vs. explicit learning may be found in Taylor and Ivry (2013).

Psychology experiments using motor learning-based methodologies to disentangle implicit vs. explicit contributions to learning have generally found a trade-off between the two: although explicit knowledge is associated with faster learning and superior performance in tasks such as visuomotor transformation (Werner & Bock, 2007), motor sequence learning (Willingham et al., 1989), and adaptation to walking on a split-belt treadmill (Malone & Bastian, 2010), explicit strategies can also hurt performance when these are applied to well-practiced sequences (Beilock & Carr, 2001; Beilock et al., 2002; Castaneda & Gray, 2007; Flegal &

Anderson, 2008; Perkins-Ceccato et al., 2003; Poolton et al., 2006; Zachry et al., 2005; Tanaka & Watanabe, 2018). To date, disagreement remains regarding the exact interplay and degree of overlap between implicit and explicit systems. For instance, according to a serial process model, implicit learning is the result of the accumulation and gradual strengthening of relatively weak stimulus-response associations, and explicit awareness allows a learner to rapidly reinforce these associations. This is exemplified by Fitts and Posner's (1967) proposed three-stage model of skill acquisition, with learners progressing through a cognitive stage (wherein verbalizable information is acquired), an associative stage (wherein routines from the cognitive stage are strengthened, with explicit attention focused on specific details of the sequence), and the autonomous stage (wherein performance reaches a ceiling once performance is fully automatic). This progression is marked by rapid initial improvements in performance followed by a phase of gradual learning wherein performance gains accrue much more slowly.

In contrast to such a serial process model with initial strategy-based performance gradually giving way to more automatized control, a parallel process model holds that implicit and explicit learning may be independent processes with no overlap or homology between them whatsoever. This framework would argue that implicit and explicit learning are qualitatively different and mutually independent throughout all stages of learning, though with a shift in the relative weight given to each of these processes as performance changes. The parallel process viewpoint is best exemplified by Logan's (1988) instance theory of automatization as well as by a more recent dual model presented by Keele et al. (2003). Note that this theoretical debate between psychological models that conceive of implicit and explicit processes as either serial, interacting processes or parallel, independent processes is similar to debates in the field of second language acquisition regarding the interface of implicit and explicit learning, as described in Section 2.3.1.

One common paradigm in the field of psychology that has contributed to debates on serial process vs. parallel process models involves a so-called serial reaction time task, in which participants press buttons corresponding to locations cued on a display (as per a methodology introduced by Nissen & Bullemer, 1987). Unknown to the participants, the cued locations may follow an underlying pattern. Learning of this pattern is indicated by faster reaction times to trials that follow the pattern than to trials that do not follow the pattern, which may be either interspersed in the main experimental block or presented in a separate block after training. An important feature of the serial reaction time task is that it can involve both implicit and explicit learning: participants may develop partial awareness of the underlying sequences and even show anticipation of the stimuli (i.e., responses that temporally precede the actual appearance of this location cues), but this explicit learning can be experimentally altered, e.g., by introducing a distracting concurrent task such as discriminating tone pitches (as in Perruchet & Amorim, 1992). Participants' explicit knowledge of the learned patterns can be assessed by post-experimental interviews wherein they are asked to recall a sequence or to indicate their predictions for the next cue location in a forced-choice task.

Previous studies using this serial reaction time paradigm have found evidence of learning of pattern sequences even when measures of explicit knowledge are near chance level, thus indicating implicit learning (e.g., Willingham et al., 1989). As further evidence that fully implicit learning can occur, learning effects on the serial reaction time task have been shown in amnesic patients with previously-demonstrated impairments in acquiring declarative knowledge (such as the famous patient H.M.; Scoville & Milner, 1957; Corkin, 1968) as well as by participants with pharmacologically-induced transient amnesia (Nissen & Bullemer, 1987; Nissen et al., 1987; Nissen, Ross, Willingham, Mackenzie, & Schachter, 1988). These findings from the serial reaction time task speak to the possibility of implicit learning, which would support findings from L2 morphosyntax learning studies as discussed in Section 2.2.

Findings from serial reaction time studies have also suggested that implicit and explicit learning co-occur with little interaction between them. For instance, in a seminal serial reaction time study that sought to gauge the relative contributions of implicit vs. explicit processes (Curran & Keele, 1993), participants with prior explicit instruction about the pattern sequence showed faster reaction times than participants who had no instruction and had to learn the sequences implicitly. However, when a distracting dual task was introduced in a later experimental phase, the explicit and implicit groups showed comparable performance. These results suggest that explicit and implicit learning systems operate in parallel—that is, simultaneously but without a strong link between the two. Under this interpretation, explicitly instructed participants used both systems under a single task condition but relied solely on an implicit system when conscious processing was hindered in the dual task condition. These seminal results have since been corroborated by other studies using different experimental tasks which have found a limited role for explicit knowledge in dual task conditions (e.g., Taylor & Thoroughman, 2007, 2008; Ewolds et al., 2017). Furthermore, neuroimaging studies have identified underlying differences in the neural activity induced by single vs. dual task conditions, as shown using photon emission tomography (Grafton et al., 1998; Hazeltine et al., 1997) as well as functional magnetic resonance imaging (fMRI) (Rauch et al., 1995; Doyon et al., 1996; Seidler et al., 2005) methodologies. In this way, experiments using a serial reaction task approach would suggest that implicit and explicit processes have a limited interface.

One possible limitation of the serial reaction time task is that it prioritizes goal-selection (i.e., of the cued location) over motor execution, in that the actual pressing of the response buttons is a relatively simple task to execute. Given that many real-world tasks—including language production—require the careful integration of planning and execution, this might limit the insights that could be gleaned from the serial reaction time task alone. By contrast, another useful experimental paradigm that involves a more intricate form of execution involves visuomotor adaptation. In such methodologies, the experimenters induce a perturbation in participants' mapping between visual and proprioceptive space, e.g., via prismatic eyeglasses (Redding & Wallace, 1988, 1993, 1997) or virtual reality systems (Cunningham, 1989; Imamizu & Shimojo, 1995; Krakauer et al., 2000). To successfully overcome these experimenter-induced perturbations, participants must learn to recalibrate an internal model of their motor system, potentially using a mixture of implicit and explicit knowledge.

One particularly common visuomotor adaptation paradigm involves an aiming task wherein participants making swiping motions on a screen must adjust for a covert rotation to successfully hit a target location (Taylor & Ivry, 2013). Before each trial, participants are asked to verbally report their intended target from a set of landmarks on the display indicating degrees (e.g., 45°, 50°, etc.). As such, the participant's reported target provides a measure of the amount of explicit re-aiming, and the difference between this reported target and the actual participant movement provides a measure of the amount of implicit (i.e., experimentally observed but unreported) re-aiming. In this way, experimenters can collect a trial-by-trial measure of implicit vs. explicit contributions to learning.

One study using this visuomotor rotation paradigm (Taylor et al., 2014) found that explicit instruction to the participants slightly increased the rate of initial learning but had only a subtle effect on implicit learning, as assessed by the reported vs. unreported components of compensatory rotation. These results were corroborated by the fact that participants who did and did not receive online task feedback both showed a comparable “after-effect” wherein implicit re-aiming persisted even after the experimentally induced perturbation was removed in a final phase of the experiment. This finding of parallel and independent operation of two learning processes supports the aforementioned findings from the serial reaction time paradigm of a dissociation between implicit and explicit learning (Curran & Keele, 1993).

The notion of implicitness and explicitness as distinct and independent components to learning is also supported by findings from the visuomotor adaptation literature that implicit learning continues even when explicit knowledge is “good enough” to meet task demands. In one study (Mazzoni & Krakauer, 2006), participants in a visuomotor rotation task received explicit instructions about the underlying perturbation in an initial experiment phase. They went on to show virtually perfect performance at early stages of testing but gradually “drifted” from the target location. This was attributed by the authors to continued implicit learning driven by non-zero prediction error between the sensory system (i.e., from participant’s actual motor systems in executing the movements) and the visual information that participants received as they performed the trials. These results align with findings of a “push-pull” dynamic between implicit and explicit learning, such that these two components may work against each other at times (Izawa & Shadmehr, 2011; Shmuelof et al., 2012). This parallels the phenomenon of blocking effects in the field of L2 acquisition, as described in Section 2.3.1 (e.g., N. Ellis, 2006; Ellis & Sagarra, 2010; Solman & Chung, 1996).

What are the differences in the actual neural processes or computations that underlie implicit and explicit learning? Insight into this question comes from computational modeling approaches that have successfully simulated the simultaneous operation of implicit-like, prediction error-based learning mechanism and an explicit, goal-based learning mechanism (e.g., Taylor & Ivry, 2011; Izawa & Shadmehr, 2011). Such research suggests that implicit learning is driven by prediction error in a "forward model" that seeks to convert an input (i.e., low-level sensations associated with aiming) to some output (i.e., target location) (Haith & Krakauer, 2013). This is suggested by findings of "drift" from the correct target in visuomotor rotation tasks even when instructed participants show initial accuracy, which suggest that the implicit system does not "have access to" the re-aiming strategy taken by the explicit system (Mazzoni & Krakauer, 2006). Instead, the mechanism that drives implicit learning simply assumes that the aiming location and the feedback location should coincide, using the disparity between the two to update some internal representation. By contrast, explicit processes rely on updating of some "inverse model" that starts with an observed output (feedback location) to deduce an input (target location) (Haith & Krakauer, 2013). Critically, the performance of explicit mechanisms may be contingent on the nature of the feedback that a learner receives (Haith & Krakauer, 2013). To illustrate this, the aforementioned visuomotor rotation study by Taylor, Krakauer, and Ivry (2014) compared the effects of presenting feedback (in the form of a cursor indicating the endpoint location) either during the trials themselves or only at the end of a trial. For participants who only received feedback at the end of a trial, gains from explicit re-aiming were higher but implicit learning was slower. Conversely, for participants who received real-time feedback, there was faster implicit learning but lower gain from consciously "exploring" other locations as targets for explicit re-aiming. As such, these findings indicate that explicit learning may differ based on the sort of information received by a learner.

Whether or not implicit and explicit mechanisms “communicate directly with each other” or “have access” to each other’s products of learning, an additional consideration is that explicit processes might influence implicit learning through indirect means. In other words, explicit processes may "bootstrap" implicit learning by allowing for the generation of strategies that are conducive to prediction error-driven learning (Taylor & Ivry, 2012). This idea of a facilitative role for explicit processes despite underlying independence between explicit vs. implicit learning mechanisms is supported by findings that spatial working memory plays a role in visuomotor rotation adaptation: in one study, participants with higher spatial working memory capacity showed faster learning, potentially because explicit mental rotation abilities compensated for low-level rotation carried out by an implicit system (Anguera et al., 2009). Similarly, age-related declines in visuomotor adaptation task performance are attenuated for participants who are able to verbally describe the underlying perturbation, when compared to age-matched adults who are unable to verbalize the perturbation (Heuer & Hegele, 2008). In this way, the outputs of an explicit system may ultimately (though indirectly) facilitate the production of the low-level task execution routines associated with a correct response, which would in turn provide the necessary inputs for an implicit system’s forward model to be updated. In this way, explicit processes may facilitate implicit learning through indirect rather than direct means. This notion is similar to the weak interface hypothesis in L2 learning as described by N. Ellis (2005).

In sum, non-linguistic studies using motor learning paradigms such as serial reaction time and visuomotor rotation tasks can be used to triangulate insights on the relative contributions of implicit vs. explicit processes to learning. As described above, some key findings are that explicit and implicit processes operate simultaneously (e.g., Curran & Keele, 1993) but involve qualitatively different underlying computations in how learning is achieved (e.g., Haith &

Krakauer, 2013). The two processes may work against each other at times, as illustrated by gradual, implicit-driven “drift” away from target locations in visuomotor rotation tasks (Mazzoni & Krakauer, 2006). Additionally, the direct effect of explicit processes on implicit learning seems to be minimal, in that implicit learning proceeds equally whether or not explicit learning co-occurs (e.g., Taylor et al., 2014). However, explicit strategies may facilitate implicit learning through indirect means, e.g., by creating the sorts of circumstances that provide the error signals necessary for implicit mechanisms to operate (Taylor & Ivry, 2012).

How can these findings from the wider field of cognitive psychology inform ongoing debates about the interface between implicit and explicit L2 knowledge? Although the findings of a limited connection between implicit and explicit knowledge (e.g., Curran & Keele, 1993) would seem to support a no-interface position (e.g., Krashen, 1981; Schwartz, 1993; Truscott, 1996; Paradis, 2009), the idea that the ultimate outputs of explicit processing routines can indirectly feed into implicit learning mechanisms (e.g., Taylor & Ivry, 2012) would seem to align with N. Ellis’s (2005) proposal of a weak interface between implicit and explicit processes, wherein conscious processing does not directly translate into implicit knowledge but rather facilitates implicit learning by inducing facilitative conditions (e.g., attentional routines, access to feedback, etc.). Furthermore, the notion that implicit and explicit knowledge can sometimes work against each other (e.g., Mazzoni & Krakauer, 2006) is captured in the notion of blocking effects in L2 acquisition (e.g., Ellis, 2006; Ellis & Sagarra, 2010; Solman & Chung, 1996).

2.3.4 Issues with generalizing findings from psychology to language processing

One important word of caution here is that findings from non-linguistic experiments may not always generalize to the acquisition of second languages. After all, models of attention from cognitive psychology have traditionally been based on research from non-SLA domains such as general pattern recognition (Posner, 1994), native language word processing (Posner et al., 1988), and (non-linguistic) semantic systems (Posner, 1992). Furthermore, non-linguistic experiments have operationalized attention and awareness only in limited contexts such as fixations towards visual locations on a display (e.g., Posner, 1995; Posner & Petersen, 1990; Posner & Rothbart, 1992). However, as Simard and Wong put it, "[o]rienting to visual locations in space has little or nothing to do with the need for learners to direct their attentional resources to features of L2 input to facilitate their intake of that input" (2001, p. 110). Put another way, to the extent that such experimental paradigms from cognitive psychology (e.g., tracking the location of a visual cue on a computer display, adjusting motor actions in response to experimenter-induced perturbations, etc.) involve only simplified low-level operations that invoke different systems of execution than more cognitively-oriented linguistic tasks, then findings from non-linguistic studies may not generalize to linguistic domains. After all, differences between the two domains are easy to think of, including for instance the limited role of perception in serial reaction time and visuomotor rotation tasks as well as the unidimensional nature of these tasks' outputs (e.g., location of a cue from a limited set of possibilities in a serial reaction time task; degree of rotation for visuomotor rotation tasks). In light of such differences between domains, the link between implicitness/explicitness in psychology and implicitness/explicitness in second language acquisition has been described as a "complicated matter, with several different but partly overlapping distinctions" (Schmidt, 2010, p. 725).

In order to overcome such issues in generalizability from non-linguistic experiments, it is critical to adopt more language-like experimental paradigms that can shed light on implicitness and explicitness specifically in the context of SLA. One such candidate paradigm comes from studies that use artificial grammars to investigate statistical learning, i.e., the attunement to probabilistic patterns in the environment (Gallistel, 1990; Kelly & Martin, 1994). In one seminal study, 8-month-old infants were found to be sensitive to three-syllable nonsense words that they were exposed to in a continuous repeating auditory stream, when compared to foil sequences wherein the same syllables were recombined in novel orders (Saffran, Aslin, & Newport, 1996). This finding of learning from artificial grammars has since been extended to older children and adults (e.g., Fiser & Aslin, 2001, 2002; Saffran, 2002; Saffran et al., 1999; Turk-Browne et al., 2005), even in the absence of instructions or conscious attempts to extract patterns and even when participants are distracted by an unrelated simultaneous task (Saffran et al., 1997; Turk-Browne et al., 2009). As such, this sort of statistical learning has been described as occurring "incidentally" (Saffran et al., 1997), "involuntarily" (Fiser & Aslin, 2001), "automatically" (Fiser & Aslin, 2002), "as a byproduct of mere exposure" (Saffran et al., 1999), and "without intent or awareness" (Turk-Browne et al., 2005)—thus fitting the definitions of implicitness as the term is used in this dissertation (for a full review of statistical learning with artificial grammars, see Romberg & Saffran, 2010).

Such a statistical learning paradigm can also be altered to specifically examine the differences between implicit and explicit learning. For instance, in one study (Batterink et al., 2015), participants were exposed to a continuous auditory stream of repeating three-syllable nonsense words, with half of the participants receiving explicit training on these words ahead of time and half of the participants being exposed to this speech stream with no prior instruction. EEG recordings and behavioral measures (namely, a speeded reaction-time task) were used to determine the extent to which participants learned these statistical regularities. The study found that participants who were explicitly trained responded more quickly to predictable targets but more slowly to less predictable targets (when compared to the no-instruction group), with the appearance of a P300 ERP component suggesting a greater recruitment of controlled, effortful processing for this group. This suggested that information is processed more automatically and less effortfully when it is learned implicitly as opposed to explicitly, though with corresponding drawbacks for performance on items that would otherwise be highly predictable when processing explicitly.

Arguing against the limits of generalizing statistical learning studies to SLA, Tomlin and Villa argue that the learning conditions and task demands of artificial grammar experiments do not match those of second language acquisition of natural languages (1994, p. 207). One particularly important critique is that artificial grammar stimuli are not *meaningful.* In other words, to the extent that language comprises a system of form-meaning connections, then the learning of forms without meaning would seem to only capture half of the picture. This limitation is overcome by Batterink and colleagues' (2014) study on rule-unaware L2 morphosyntax acquisition, as described in section 2.2.5. Namely, given that the novel articles used in this experiment (i.e., *gi*, *ul*, *ro*, and *ne*) actually encoded a meaning of some sort (specifically, the distance to the referent in the overt rule, and the animacy of the referent in the covert rule), then this experimental paradigm parallels the learning task that learners face when trying to acquire any morpheme in a natural language. As such, Batterink et al.'s (2014) experimental design presents a good "test-tube model" of language learning (to borrow a phrase from Ettlinger et al., 2016, p. 825). This is particularly true in light of prior research showing that performance on artificial language learning is positively correlated with classroom-based measures of second language learning ability even when controlling for general cognitive factors like IQ (Ettlinger et al., 2016), thus supporting the validity of artificial languages as an experimental paradigm in L2 psycholinguistics research.

Although Batterink et al.'s (2014) experiment speaks more directly to actual second language learning than do non-linguistic studies or artificial grammar studies (thanks to its use of meaningful language stimuli), and although its findings of neural correlates of implicit and explicit processing lend support to previous behavioral work on rule-unaware learning (e.g., Leung & Williams, 2012), several limitations in its study design prevent us from making certain inferences about the interplay of implicit vs. explicit processing. In particular, its univariate ERP-based analysis method means that neural markers of implicit vs. explicit processes may have obscured each other to some extent, making them hard to detect if they occurred simultaneously. Furthermore, the study design could not speak to the possible role of prediction as a strategy that distinguishes rule-aware from rule-unaware learning. Finally, the relationship between behavioral reaction times and ERP components (which might indicate whether conscious processes feed into task performance in-rule aware learners) remains to be explored. Each of these issues is further described and addressed in the extensions described below.

2.3.5 Motivation for extensions of Batterink et al. (2014)

Extension 1: Separately quantify EEG indices of implicit and explicit processing

One limitation to extrapolating Batterink et al. (2014)'s findings to the interplay between implicit and explicit processing is that the study design could not determine whether reaction-time measures of learning reflected underlyingly similar or different processes in the rule-aware vs. rule-unaware participants. Correlations from sleep data (which was collected between two training phases) suggested that a similar underlying neural mechanism was involved between rule-aware and rule-unaware participants, because sleep had similar effects of improving learning for both groups. However, the ERP data could not speak to this fully: although separate brain activation patterns were found for participants who did and did not report being aware of the hidden regularity (with rule-unaware participants showing an early negativity and rule-aware participants showing a P600 component), time window overlap issues between these two components meant that these neural indices could not be quantified separately in a univariate analysis. In other words, because of potential temporal overlap between the two ERP components of interest (the early negativity and the P600), a univariate measure that only reflects the sum of these two voltage deflections at any given time period would be unable to determine whether and to what extent implicit processing of the rule violation was occurring in the rule-aware participants. Note that similar issues caused by temporal overlap also limit the interpretations of the previously mentioned EEG experiments that used implicit vs. explicit training conditions (Morgan-Short et al., 2012) and subjective measures of source attribution (Wan et al., 2010).

Thankfully, these limitations with disentangling neural markers associated with implicit and explicit L2 morphosyntax processing may be addressed by using Multivariate Pattern Analysis (MVPA). This methodology builds off of recent developments in the field of machine learning, advances which have been incredibly fruitful for research fields as disparate as genetics (e.g., Libbrecht & Noble, 2015), cancer treatment (e.g., Kourou et al., 2015), drug discovery (e.g., Lavecchia, 2015), animal behavior (e.g., Valletta et al., 2017), energy conservation (e.g., Marasco & Kontokosta, 2016), and earthquake prediction (e.g., Li et al., 2018). These artificial intelligence-based methods are now common in fMRI research for decoding brain activity patterns that are impossible to identify solely through traditional univariate analyses (Haxby et al., 2001; Haynes & Rees, 2006), but its extension to EEG analysis has come only recently (King & Dehaene, 2014; Grootswagers et al., 2017). MVPA has been demonstrated to be able to detect effects that cannot be found by traditional ERP averaging of univariate voltages alone (e.g., Cauchoix et al., 2012, 2014; Fahrenfort et al., 2017). As one prominent EEG researcher wrote, "In my research so far, I have been able to detect very subtle effects that never would have been statistically significant in conventional analyses" (Luck, 2018).

In MVPA, machine learning algorithms are trained to predict the experimental condition to which trials belong by learning to detect multivariate neural activation patterns associated with those conditions. This method is *multivariate* in that the relationships between different variables (in this case, voltage readings at separate electrode channels in the EEG recording) are taken into account, as opposed to univariate measures (like event related potentials) that treat readings from each electrode as essentially independent. MVPA works by representing each trial as an individual datapoint in a multidimensional space wherein each dimension represents voltage in a separate electrode. A machine learning algorithm then finds the boundaries that best divide datapoints (i.e., trials) from different experimental conditions. This process is performed separately for each time point in the EEG epoch window. If MVPA classifiers trained to detect such patterns within a given dataset can achieve high accuracy when classifying trials from a different dataset, then one can infer substantive similarity of the underlying neural activity across the two datasets (Kaplan, Man, & Greening, 2010).

MVPA has several advantages over traditional univariate EEG analysis methods. For instance, because MVPA classification is performed on individual trials, MVPA can be used to track the degree to which different neurocognitive events share neural patterns at the level of individually classified trials, in contrast to traditional null hypothesis testing of trial-averaged ERP patterns (Grootswagers et al., 2017). Furthermore, because MVPA automatically extracts information relevant to the difference between experimental conditions rather than requiring experimenter specification of analysis parameters (e.g., about the specific electrodes to analyze), this method can detect experimental differences that would be much harder to identify or substantiate through univariate approaches alone when the nature of the effect is unknown *pre hoc* (Kriegeskorte et al., 2009; Fahrenfort et al., 2018). As an additional advantage, MVPA decoding is typically performed on a per-participant level, which helps to account for individual differences in language processing signatures (e.g., Tanner & Van Hell, 2014) and thus to avoid issues from traditional ERP analyses that aggregate data from different participants and analyze them identically with the same parameters (i.e., using the same time windows and electrode channels).

The most important advantage of MVPA for the purposes of this dissertation is that it can be used to disassociate neural components even when these overlap in time. This is done by gauging the extent to which an MVPA decoder that is trained to classify trials based on data from one time point can generalize to other time points (i.e., show above-chance accuracy when tested on data from *other* time points in the EEG trial). In this way, MVPA can measure distinct neural patterns even when these patterns overlap temporally, thus potentially canceling each other out in a simple univariate EEG signal. This is precisely the case with the EEG experiment presented in Batterink et al. (2014). In fact, the authors specifically identify this problem in their discussion section, writing:

> It is possible that Rule-Aware participants also acquired implicit knowledge, but that the negativity did not reach statistical significance in this group because of overlap from the early portion of the P600 effect (p. 177).

As a concrete illustration of how temporally overlapping activity can be disentangled using MVPA, Heikel et al. (2018a) disassociated the N400 and P600 ERP components as follows. They first trained separate MVPA decoders: one that was trained to detect neural activity from the N400 time window, and another that was trained to detect neural activity from the P600 time window. They then separately tested the performance of these two decoders on other time points across the EEG trial. The authors found that each of these decoders showed above-chance accuracy even when tested on time periods outside of the time windows that they were originally trained on. This indicated that the associated activity for each of the decoders did not occur only in the time window associated with their respective ERP component, but also occurred before and after those time windows to some degree. Furthermore, the time periods of above-chance accuracy for each of the decoders overlapped to some degree, thus suggesting that the neural activity associated with the N400 and the P600 overlap temporally. Each of these decoders showed separate distinct peaks of maximum trial classification performance for different time points in the EEG trial. Critically, when the decoders' performances were compared to each other time point by time point, they were found to perform significantly better/worse than each other at certain points in the trial, thus suggesting that the decoders were indeed tuned to detect different kinds of neural activity. These MVPA results suggested that previously reported ERP patterns from univariate analyses were not the result of a single process that was shifted in latency across experimental conditions, but rather the result of two different components that overlapped temporally to some degree, as had been hypothesized in previous literature (Bornkessel-Schlesewsky & Schlesewsky, 2009; Kuperberg, 2007).

For my first extension of Batterink et al. (2014), I train an MVPA decoder to detect neural activity associated with implicit (i.e., rule-unaware) processing of rule-adhering vs. rule-violating trials in the context of an artificial language experiment (i.e., to detect the neural activity that was manifested as an early negative ERP component in Batterink et al., 2014) and subsequently determine whether this same decoder can accurately distinguish rule-adhering from rule-violating trials when participants are aware of the underlying rule. If such a decoder achieves above-chance classification accuracy, then this would suggest that participants who are aware of a morphosyntactic rule are also employing some of the same neural mechanisms as when they are unaware of the rule. Conversely, if the decoder cannot achieve above-chance accuracy, then entirely distinct processes would seem to be involved. Stated formally, the research question is as follows:

Research Question 2: Do rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data?

Extension 2: Measure prediction of semantic features

Another way that MVPA can be used to disentangle implicit vs. explicit processing is by examining neural indices of real-time prediction in morphosyntax processing to determine if this is a feature that distinguishes rule-aware vs. rule-unaware processing. This builds from a prominent line of research within the field of L2 acquisition that examines prediction for target linguistic forms in the upcoming input during real-time language processing. As an example, some studies suggest that native Spanish speakers (but not Spanish L2 learners) use morphosyntactic cues from grammatical gender to facilitate processing of upcoming nouns in the speech stream (Lew-Williams & Fernald, 2007). Such findings of linguistic prediction have been subsequently corroborated using methods such as behavioral response times (e.g., Federmeier et al., 2010), eye-movement tracking (Lew-Williams & Fernald, 2007), and EEG (e.g., Lau et al., 2013). Note that, in their discussion section, Batterink et al. (2014) overtly identify as an open question the possibility of differences in prediction between rule-aware and rule-unaware participants. The authors write:

> … Rule-Aware participants may have adopted a different strategy for processing the stimuli, perhaps forming conscious expectations of the article noun pairings. Thus, a similar RT [Reaction Time] delay may reflect implicit

> learning in Rule-Unaware participants, and strategic, explicit processing in Rule-Aware participants. (p. 177).

To illustrate how MVPA can be used to examine prediction, one recent study on native speaker listening (Heikel et al., 2018b) trained an MVPA decoder on EEG responses to animate vs. inanimate nouns. In subsequent testing, this decoder showed above-chance accuracy in sorting animate vs. inanimate trials when tested on the silent periods immediately before the presentation of a noun whose (in)animacy was predictable given the preceding sentence context. These results indicate that, at least in native language listening, semantic features like (in)animacy are activated in anticipation of animate and inanimate nouns. Importantly for my dissertation, this study provides a proof-of-concept for using MVPA to study prediction in language processing.

In order to extend the study design of Batterink et al. (2014) to examine linguistic prediction, one could train an MVPA decoder to distinguish participants' neural processing of animate vs. inanimate nouns (based on EEG data collected when they read English nouns alone in a pre-task block), and subsequently test this animacy-sensitive decoder on their neural activity when processing the novel artificial language articles in isolation during the main experiment task. In this way, one could determine whether the same sort of neural activity that distinguishes their reading of animate vs. inanimate nouns (e.g., *cat*, *dog* vs. *table*, *chair*) also occurs when participants are reading the novel articles that are predictive of animacy vs. inanimacy (e.g., *gi*, *ul* vs. *ro*, *ne*). Running this analysis individually on either the rule-aware or rule-unaware group allows one to see whether prediction occurs within that group; direct t-test comparisons across groups allow us to see whether the two groups differ significantly in prediction. One possibility is that rule-aware participants begin actively anticipating (in)animate entities when they process the novel articles, suggesting that any performance differences between the rule-aware vs. rule-unaware groups may be partly driven by conscious predictive strategies. Conversely, another possibility is that rule-aware and rule-unaware participants do not differ in their real-time semantic prediction, either because neither or both groups use prediction.

These results may lend insight into explaining variability in L2 learning success that might stem from different learners' processing styles. For instance, prediction may be a useful strategy for language processing if it helps learners by reducing the number of possible upcoming linguistic forms that they must anticipate in the upcoming L2 input stream. This seems particularly likely given findings of lexical competition in real-time language processing, both for native language (e.g., Norris et al., 1995; Brouwer & Bradlow, 2016) and second languages (e.g., Weber & Cutler, 2004; Broersma & Cutler, 2011). To illustrate this in the context of my experiment, a participant who predicts an animate noun after reading an artificial language article might limit their pool of possible word candidates to animate words only. Consequently, they might show better performance in processing the upcoming noun and determining that it is a living entity, when compared to a participant who anticipates any noun in general and would thus have many more possible candidate word forms to "prepare for."

This use of MVPA would make a novel contribution to research on L2 linguistic prediction by directly examining its underlying neural implementation. To date, research on L2 prediction has only used univariate measures of prediction such as response time slow-downs (e.g., Federmeier et al., 2010), eye-movement behavior (e.g., Lew-Williams & Fernald, 2007), and ERP indices of general surprisal as opposed to access of specific meanings (e.g., Lau et al., 2013). Experiment designs that use such limited measures must be carefully constructed to study participants' responses to a limited set of comparison conditions. Furthermore, because such studies use as their dependent variable a downstream consequence of prediction rather than measures of the cognitive process of linguistic prediction itself, such experiments can only determine that one linguistic form was more/less predictable than another one, and not that a certain linguistic form was specifically predicted for. As such, experiments on L2 prediction to date have used experimental stimuli that may not reflect the actual target of linguistic prediction (e.g., if they examine differences between more/less predictable items but might not include the actual predicted items as an experimental stimulus) and have used dependent variables that do not measure prediction directly, in the way that it is implemented in real time. These issues can be overcome by an MVPA approach wherein the training data for the decoder comprises the actual neural activity that encodes the target of prediction, and the testing data for the decoder comprises the neural activity from the time period when prediction might occur, as opposed to neural activity that only occurs later as an indirect consequence of prediction. Stated formally, the research question is as follows:

Research Question 3: Do rule-aware and rule-unaware participants show neural evidence of semantic prediction from a covert morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so, do they differ in their use of prediction?

Extension 3: Examine relationship between explicit processing and behavioral responses

Another feature that may distinguish implicit and explicit language processes is their relationship with externally observable performance in a linguistic task. Although Batterink et al. (2014) found no group-level reaction time differences between the rule-aware and rule-unaware groups, the study could not determine whether the reaction time measures in the behavioral task were quantifying the same underlying process between the two groups—in other words, whether reaction-time effects were tied to implicit or explicit processes. The authors specifically identify this limitation in their discussion section, writing:

> The RLI [Response Learning Index behavioral measure of learning] did not vary as a function of rule awareness. This result may be interpreted in at least two different ways. The first possibility is that the RLI represents an implicit index of learning, occurring independently of rule awareness... An alternative interpretation to consider is that although the RLI is similar between the two groups, it reflects different underlying causes. (Batterink et al., 2014, p. 177)

This limitation can be addressed through a fine-grained analysis of rule-aware participants' P600 ERP component latency, which would allow us to determine the contributions of implicit vs. explicit processing to externally observable language-related behavior (measured via reaction times in an online linguistic task) as described below.

Standard practice in L2 psycholinguistics is to analyze the timing of the P600 ERP component relative to the onset of stimulus presentation rather than relative to an external task-related response (e.g., Friederici et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992). Perhaps as a result of this, the (presentation-locked) latency of the P600 component has been found to vary, both between participants (Kotz, 2009) and between experimental conditions (e.g., Allen et al., 2003). As an extreme illustration of variability in the timing of this component, some studies have found P600 effects in entirely separate, non-overlapping time windows, e.g., 500-800 ms post-stimulus (Kalatzis et al., 2004) vs. 800-1300 ms post-stimulus (Rossi et al., 2005). However, a recent study (Sassenhagen et al., 2014) suggests that the P600 does not vary in latency when its onset is defined relative to the timing of participant's responses rather than to the timing of stimulus presentation. This response-locked nature of the P600 response was found to be shared with the P3b component (which has been associated with conscious processing, Rutiku & Bachmann, 2017) but not with the N400 component (which has been interpreted as a stable marker of automatic processing, e.g., because it has been detected even when word stimuli are presented subliminally and when participants cannot report the meanings of the presented words, Luck, Vogel, & Shapiro, 1996). These findings suggest that the P600's appearance in

Sassenhagen, Schlesewsky, and Bornkessel-Schlesewsky's (2014) experiment was driven by a conscious response to the eliciting word rather to more automatic, potentially subconscious (i.e., implicit) processes.

In this way, a response-locked approach to analyzing the P600 component goes beyond the inherent properties of the eliciting linguistic stimuli themselves (which might include factors like grammatical correctness, semantic plausibility, etc.) and instead takes a closer look at the role that the P600 component plays in the causal chain of events that begins with stimulus presentation and perception, proceeds to conscious stimulus processing (as indexed by the P600), and ends with an external behavioral response. This interpretation of the P600 component as an index of conscious processing that feeds into external task performance might help to explain previous findings regarding variability in P600 latency. For instance, findings of a delayed P600 to less frequent words (Allen et al., 2003) may be attributed to delays in consciously processing rarer words that require more time to access. Similarly, delayed P600 components when processing one's second language (Kotz, 2009) might be attributed to difficulties in processing vocabulary or grammar that is less familiar or less readily accessible. These findings of delayed P600 components in adverse language processing conditions run parallel to findings that the (also consciousness-related) P3b component is similarly delayed when working memory is burdened (e.g., Naccache et al., 2016).

For the purposes of this dissertation, one could implement a fine-grained analysis of ERP component latency to determine whether rule-aware participants' P600 components to rule-violating stimuli are time-locked to overt behavioral responses—in which case the P600 could be interpreted as a prerequisite for response—or not time-locked—in which case the P600 would seem to be functionally unrelated to the response. In turn, this would allow us to infer whether reaction times in the artificial language learning task are driven by explicit as opposed to implicit processes in rule-aware participants. This line of inquiry is highly relevant to second language teaching pedagogy, to the extent that an L2 educator's ultimate goal is to improve learners' ability to carry out observable linguistic behaviors in their external environment (whether through implicit or explicit processes). Stated formally, the research question is as follows:

Research Question 4: Is the P600 component of explicit processing time-locked to rule-aware participants' external behavioral responses in an artificial language learning task?

2.4 Overview

In sum, the overall aim of this dissertation is to provide further insight into the interplay between implicit and explicit processes in second language morphosyntax acquisition. This topic of research is informed generally by disagreements in SLA theory regarding whether L2 acquisition can occur without conscious awareness of form-meaning connections, as proposed by some models (Tomlin & Villa, 1994; Leow, 2015) but explicitly denied by others (Schmidt, 1990; Robinson, 1995). This line of inquiry also speaks to theoretical debates surrounding the interface between implicit vs. explicit processes in second language acquisition, with competing models arguing for no interface (wherein all learning is implicit in nature; e.g., Paradis, 2009), a strong interface (wherein implicit and explicit processes are isomorphic; e.g., DeKeyser, 2007), or a weak interface (wherein most learning is implicit but consciousness plays a facilitative role; e.g., Ellis, 2006).

Although previous behavioral studies in the field of SLA have found evidence for learning in both explicit and implicit *conditions* (Norris & Ortega, 2000; Spada & Tomita, 2010; Goo et al., 2015), these only speak to the training environment rather than to L2 learning at the level of individual learners. In turn, experiments analyzing rule-aware vs. rule-unaware learning solely through behavioral methodologies (e.g., Williams, 2004, 2005; Hama & Leow, 2010; Faretta-Stutenberg & Morgan-Short, 2011; Leung & Williams, 2012, 2014; Rebuschat et al., 2013) are limited in their ability to determine the exact nature of processing occurring in learners over the course of the experiment. Furthermore, although previous studies from the wider field of cognitive psychology have examined implicit vs. explicit learning using a variety of methodologies like serial reaction time paradigms (e.g., Curran & Keele, 1993), visuomotor tasks (e.g., Anguera et al., 2009), and artificial grammars (e.g., Batterink et al., 2015), the fact that these studies do not involve learning form-meaning connections brings up the possibility that their results may not be generalizable to the domain of language processing (Morgan-Short, 2020). As such, experiments that combine artificial language learning paradigms (as in Leung & Williams, 2012) with EEG methodologies (as in Batterink et al., 2014) may provide the best approach for exploring the role of implicit vs. explicit processes in L2 morphosyntax acquisition.

2.4.1 Summary of my study

This dissertation aims to reproduce Batterink et al. (2014) and to implement EEG analysis techniques that are novel to the field of L2 psycholinguistics in order to address three open questions that were specifically identified in the original study, regarding the possible co-occurrence of implicit and explicit processing; the potential use of different strategies in rule-aware vs. rule-unaware learners; and the nature of the underlying processes captured by behavioral response time measures. The reproduction component seeks to determine whether rule-aware and rule-unaware participants show differences in learning of a covert morphosyntactic animacy rule in an artificial language, as indexed by ERP components. The first extension implements multivariate pattern analysis (MVPA)—a novel method that can pull apart neural activity tied to language-related ERP components even when these overlap time (Heikel, Sassenhagen, & Fiebach, 2018a)—to separately quantify the neural markers of implicit and explicit processing. The second extension tests the possibility that rule-aware and rule-unaware participants use different strategies—i.e., predicting an (in)animate entity after reading the artificial language articles—by using MVPA to measure the extent to which neural activity tied to animacy processing can be detected prior to the perception of (in)animate nouns in the experiment (as in Heikel et al., 2018b). Finally, the third extension investigates the relationship between explicit processing and linguistic task performance by examining whether rule-aware participants' P600 components are temporally related to their reaction times, which would suggest that conscious reactions are a prerequisite for behavioral responses (as in Sassenhagen et al., 2014); this would provide valuable insight into the underlying processes captured by response time-based paradigms. For convenience, the research questions are formally restated below:

Research Question 1

Can event-related potentials capture differences in learning of a covert morphosyntactic regularity between rule-aware and rule-unaware learners?

Research Question 2

Do rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data?

Research Question 3

Do rule-aware and rule-unaware participants show neural evidence of semantic prediction from a covert morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so do they differ in their use of prediction?

Research Question 4

Is the P600 component of explicit processing time-locked to rule-aware participants' external behavioral responses in an artificial language learning task?

CHAPTER III. MATERIALS AND METHODS

This section presents the study design (Section 3.1) and analysis methods (Section 3.2) for the dissertation study. Except where otherwise indicated, this study aims to reproduce the procedure used by Batterink et al. (2014) as faithfully as possible. The overall structure of the study including its trial structure is given in Figure 2.

The experiment began with a noun-only block to allow for decoding of noun animacy/inanimacy and then a pre-training session introducing the four artificial language articles. This was followed by a main reaction time-based experimental task with simultaneous collection of EEG data. Subsequently, a structured debriefing interview was implemented to assess participant awareness of the artificial language's underlying morphosyntactic rule. Finally, a third block of the experiment was performed, after participants had been explicitly instructed about the underlying pattern. Behavioral and univariate ERP analyses were conducted to attempt to replicate the effects described in Batterink et al. (2014), as per Research Question 1. Multivariate Pattern Analyses were conducted to address Research Questions 2 and 3—namely, to separately quantify neural indices of implicit vs. explicit processing as well as to detect neural activity related to linguistic prediction induced by reading of the novel pseudoword article. Finally, reaction time-locking analyses were conducted to analyze the relationship between behavioral task responses and the timing of the P600 component, as per Research Question 4.

3.1 Participants

Participants for this study were 52 right-handed, neurologically typically developing native speakers of English with normal or corrected-to-normal vision (32 female, 20 male, *M* age = 18.71 years, *SD* = 1.29 range: 18-24). Handedness was assessed via the standardized Edinburgh Handedness Inventory (Oldfield, 1971). No participants reported hearing problems, language disabilities, learning disabilities, or a history of head trauma. All participants were undergraduate students at a large university recruited through introductory psychology courses. Participants provided consent as per institutional review board standards and receive research participation credit for their psychology course for completing the study. Demographic information about the participants is given in Table II. Of these 52 participants, 36 (69.23%) reported proficiency in a language beyond English.

For this study I chose to include native English speakers regardless of experience with second languages. Although prior experience with languages beyond English has been tied to differences in performance in previous research employing an artificial language paradigm (Williams, 2004, Experiment 1; 2005), my reasons for not controlling this factor are fourfold. Firstly, learning effects have also been reported for sample participant groups that have a mixture of different L2 experience levels (e.g., Leung & Williams, 2012). Secondly, Batterink et al.'s (2014) sole language-related criterion was that participants be native speakers of English, and so following this criterion would lead to a more faithful reproduction of their study. Thirdly, this more relaxed criterion would maximize the sample size for the experiment. Fourthly, the effects of second language experience were not the main question of interest for this dissertation. These reasons motivated my decision to use native proficiency in English as the only language background-related criterion for this study.

Nevertheless, in order to provide more in-depth participant demographic data for the purposes of future exploratory analyses, at the beginning of the study participants completed a background questionnaire. This covered information regarding age, gender, highest level of education, undergraduate major, and listing of languages known to the participant as well as their environment of exposure (i.e., home, school, and/or other) and Likert scale ratings of listening, speaking, and reading abilities for each of these languages, among other questions. This was based on the standardized LEAP-Q (*Language Experience And Proficiency Questionnaire*; Marian et al., 2007) and can be found in Appendix A.

3.2 Artificial language stimuli

Building off of the methodology used by Batterink et al. (2014) (which was in turn based on Leung & Williams, 2012), participants in my study were trained on an artificial language comprising four novel articles: *gi*, *ro, ul* and *ne*. Participants were instructed that these articles function like the English word “the” but also carry a meaning of relative distance. Namely, two of these articles (*gi* and *ro*) denote a referent that is nearby and the other two (*ul* and *ne*) denote a referent that is far away. Critically, these articles followed a regularity in that they predict the animacy of the co-occurring noun (as illustrated in Panel B of Figure 2). As shown in Table I, two of the articles co-occurred more frequently (namely, in 6/7 of trials) with animate nouns, and the other two co-occurred more frequently (in 6/7 of trials) with inanimate nouns. Participants were not told about this underlying regularity.

3.3 Procedure

3.3.1 Noun-only block

In the first phase of the experiment, participants saw English nouns presented in isolation and responded whether these nouns referred to animate or inanimate referents. This was done in order to be able to train an MVPA decoder to distinguish neural activity tied to the animate/inanimate distinction without “contamination” of the data from having seen the preceding *ul/gi/ro/ne* artificial language article. To illustrate this, the animate nouns in my experiment were (by design) preceded by *gi* and *ul* in the majority of trials, whereas the inanimate nouns were preceded by *ro* and *ne* in the majority of trials; this means that neural data from noun-reading in the main experimental task (see below) could also reflect residual effects of having just read *gi*/*ul* vs. *ro*/*ne*, meaning that this data would be unsuitable for training an animate/inanimate decoder because the decoder could inadvertently become attuned to the distinction between *gi*-/*ul*-preceded trials and *ro*-/*ne*-preceded trials. As such, in an initial phase of the experiment I collected data from English nouns presented in isolation.

Each trial began with a fixation cross (presented for 1000 ms) followed by a noun (presented for 1000 ms) and a blank screen (for a randomly selected duration between 500 ms and 3000 ms) before a prompt for the animacy response appeared. Response buttons were randomized on each trial so that for some trials the J key represented a “living” response and the K key represented a “nonliving” response whereas the reverse was true for other trials. The response assignment was communicated to participants with a response prompt that either said “living” on a box to the left and “non-living” on a box to the right or vice versa. Making participants wait before providing a button press; jittering the waiting period; and randomizing the button responses on each trial were all features designed so that participants would have the livingness/nonlivingness of the noun in mind with minimal interference from anticipating a motor response in this task, thus ensuring that my MVPA decoder for the animate/inanimate distinction was not inadvertently capturing the differences between pressing the J vs. K key.

3.3.2 Vocabulary pre-training

In the next phase of the study, participants performed a vocabulary pre-training to introduce them to the novel articles used in the experiment (*ul*, *gi*, *ro, ne*). Participants were instructed only on the near/far meaning of the articles. The pre-training itself comprised the following:

Initial instruction. Participants saw the following written explanation (the full instructions for the experimental tasks are given in Appendix B):

> In some languages, the distance of an object being referred to is reflected in the grammar. In this experiment, 'gi' and 'ro' are used with objects that are 'near,' and 'ne' and 'ul' are used with objects that are 'far.' For example, the watch that you are wearing could be referred to as 'ro watch', which would mean something like 'the-near watch.' Thus, these words combine the English meaning of 'the' with the meanings of 'near' and 'far.'

Forward translation. For this first computerized task, participants were prompted with a written English definition, i.e., *near* or *far*, and had to press a keyboard button corresponding to either *gi*, *ro*, *ul*, or *ne*. The assignment between the near-denoting words *gi* and *ro* and between the far-denoting words *ul* and *ne* was indicated via different font colors as in the initial flashcard study phase. Participants performed 24 such trials. As in Batterink et al. (2014), if on the last 12 of these trials they did not get 11 trials correct, then they performed another sequence of 12 trials until they reached this criterion.

Backward translation. Following the forward translation task, participants were prompted with the spoken novel articles *gi*, *ro, ul*, or *ne* and asked to press a button corresponding to the English definition (*near* or *far*). Each participant performed 48 trials of this task.[9]

3.3.3 Main experimental task

After the pre-training was complete, participants proceeded to the main experimental task. The overall trial structure is illustrated in Panel A of Figure 2. Each trial consisted of the presentation of a novel article (*ul*, *gi*, *ro*, or *ne*) followed by a noun. Half of these nouns were animate (e.g., *horse*, *puppy*) and the other half were inanimate (e.g., *table*, *kettle*). The

[9] Although no criterion was applied for this part of the vocabulary pre-training (using the same procedure as in Batterink et al., 2014), *post hoc* analyses on data from a behavioral pilot study showed that participants had high accuracy on this task (mean = 90.16%, standard deviation = 8.81%) that was significantly above chance levels, $t(23)= 22.33$, $p < .001$.

relationship between the articles and the animacy of the following noun was probabilistic, following the design in Batterink et al. (2014) and mirroring patterns found in natural languages. Namely, six out of every seven trials were rule-adhering rule-adhering in that they conformed to the animacy assignment presented in Table I, with *gi* and *ul* preceding animate nouns and *ro* and *ne* preceding inanimate nouns. Meanwhile, one randomly selected (violation) trial in each set of seven consecutive trials had the opposite assignment.

Participants were tasked with providing two speeded responses to each trial: whether the noun was living or non-living, and whether it was near or far (based on the preceding artificial language article). Each experimental train began with the presentation of a fixation cross (+) for 1000 milliseconds, followed by a novel article for 350 milliseconds and finally a noun for 500 milliseconds or until the animacy response is provided. In cases of an animacy response given after the 500 ms time window, a blank screen replaced the noun on the display. After the animacy response, participants were shown the cue "*Near/Far?"* until the second (near/far) response was provided. This timing followed the trial design in Batterink et al. (2014) as closely as possible, except for the 350 ms timing for the pseudoword article which was shown separately in our staggered design. Because the pseudoword could only be one of four possible two-letter options, a shorter 350 ms duration was chosen to keep the overall experiment length relatively short. Again as in Batterink et al. (2014), four responses on a standard keyboard were configured so that the four unique responses (living/nonliving/near/far) each had unique assigned buttons. Over the course of this main task, participants performed a short initial practice block of six (rule-adhering) trials followed by three learning blocks. Each learning block consisted of four (rule-adhering) buffer trials followed by 308 experimental trials (a mixture of 264 rule-adhering trials and 44 violation trials). A timed break with a fixed five-minute duration was given between the first two blocks to attenuate participant fatigue, at which point the participant was offered water and snacks. Between the second and third blocks of the experiment a rule awareness debriefing interview was conducted (see below). The stimuli were presented on a computer monitor about 100cm in front of the participant, and the experiment was coded in PsychoPy® software (Version 2020.2.3; Peirce, 2007).

As in Batterink et al. (2014), learning of the covert animacy rule was determined by comparing response times and accuracies to violation trials vs. rule-adhering trials. To avoid any confounds related to the specific nouns assigned to the rule-adhering or violation conditions, stimuli were counterbalanced within cycles of seven participants in such a way that a given noun was presented in the context of a rule-adhering trial for six out of seven participants and in the context of a violation trial for the seventh participant. Assignment of trials to either the first, second, or third experimental block was also counterbalanced across participants. Additionally, for each participant, the nouns assigned to rule-adhering and violation conditions were matched on a group level for orthographic word length (range of mean word length = 5.8-6.3 letters) as well as for frequency in the English language as per the Kucera-Francis database (Kucera & Francis, 1967); concreteness, or the degree to which a word refers to a perceptible entity, as per the normed 40,000-word database in Brysbaert et al. (2014); valence, or the extent to which a stimulus is emotionally positive or negative, as per the database in Warriner et al. (2013); and arousal, or the stimulus's degree of physiological activation, i.e., how calming or exciting/agitating a stimulus is, as per Warriner et al. (2013). These factors were also balanced across words for living vs. nonliving things. The full list of nouns used in this study is given in Appendix C.

Note five differences between this procedure and that of Batterink et al. (2014). Firstly, the article was presented separately from the noun (rather than presenting both words simultaneously) so that neural activity induced by the artificial language article could be analyzed in isolation as per Research Question 3. Secondly, because of logistical reasons and because sleep is not the primary research topic for this study, the 90-minute nap session between the two blocks was replaced by a five-minute break. Thirdly, I collected a third block of experiment data subsequent to the rule awareness debriefing rather than just two blocks of data, allowing us to collect both rule-aware and rule-unaware data from the same participants. Fourthly, to allow for a third block of trials in the experiment as well as a noun-only block, and in order to balance word features like concreteness and imageability across animate/inanimate words (so as to allow for confound-free decoding of an English noun's living or nonliving status in Research Question 2), I expanded my word stimuli for the living category beyond animals to also include arguably more abstract entities like humans (with words like "professor," "student," etc.). The full list of English nouns is presented in Appendix C. Finally, we added a "noun-only" block to the beginning of the experiment to be able to train an MVPA decoder to distinguish neural responses to living/nonliving nouns. Although this was introduced at the beginning of the experiment before the artificial language articles were first taught (such that in principle the sequence of experiment phases from Batterink et al. [2014] was left intact), this contributed to the overall length of the experiment and thus potentially to participant fatigue.

3.3.4 EEG recording and preprocessing

For the EEG sessions, participants sat in a comfortable chair approximately 100 centimeters from a 12-inch monitor in a sound-attenuated EEG recording booth. Continuous recording was performed in DC mode, with a sampling rate of 512 Hz from 32 Ag/AgCl electrodes embedded in an ANT Neuro-brand Waveguard™ elastic cap and distributed in standard and extended 10-20 locations (Jasper, 1958: FP1, FPz, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, POz, O1, Oz, O2), as well as from electrodes placed on the left (M1) and right (M2) mastoids. After applying the cap, the experimenter used electrolyte gel (pumped in via a blunt-tip needle) as well as gentle abrasion with a small wooden stick to lower all impedances below 5 kΩ. Scalp electrodes were referenced online to the common average of all electrodes. A vertical electrooculogram (VEOG) was recorded with two electrodes placed above and below the right eye, and a horizontal electrooculogram (HEOG) was recorded with two electrodes placed on the right and left canthi. As a further measure to avoid muscle-movement-related artifacts, participants were instructed to minimize eye and body movements during item presentation. The EEG signal was amplified at 22 bits using an ANT Neuro bioamplifier system (AMP-TRF40AB Refa-8 amplifier), digitized with a 512 Hz sampling rate, and filtered with a digital finite impulse response low-pass filter with a 138.24 Hz cutoff (sampling rate * .27), as implemented by ANT Neuro's asa™ recording software.

All offline data processing was carried out using the EEGLab (Delorme & Makeig, 2004) and ERPLab (Lopez-Calderon & Luck, 2014) MATLAB toolboxes and informed by the procedures from Kappenman, Farrens, Zhang, Stewart, and Luck (2021). First, data were downsampled to 256Hz and re-referenced offline to a mean of the mastoid electrodes (M1 and M2). Then, DC offsets were removed and data were filtered offline with a Parks-McClellan Notch filter set at 60 Hz (180-order) to remove line noise data, and then subsequently by an IIR-Butterworth band-pass filter (DC set to high-pass 0.1Hz, low-pass 30.0Hz based on a 200 ms moving window with a 50 ms moving step, with 12dB/octave and 40dB/decade attenuation).

Next, for each participant independent component analysis (ICA) was performed and artifacts were automatically corrected using the automatic routine in the ICLabel toolbox (Pion-Tonachini et al., 2019). Epochs were extracted from -200 to 1200 ms relative to the onset of the presentation of the noun (as in Batterink et al., 2014) as well as from -200 to 2000 ms relative to the onset of the presentation of the artificial language article (to accommodate slower reaction times in my Research Question 4 analysis). Subsequently, I performed the artifact detection routines from Kappenman et al. (2020), which rejected a) trials with extreme values with amplitudes greater than 200 µV or lower than -200 µV in any electrode; (b) trials with significant drift in the scalp electrodes with peak-to-peak thresholds exceeding 160 µV (window size = 500 ms window, step size = 100 ms), and (c) trials with horizontal eye movements in the ICA-corrected HEOG with step-like artifacts exceeding 64 µV (window size = 100 ms window, step size = 10 ms step). In addition, to reject trials that contained eye movements that significantly overlapped with the presentation of the target word, I also removed (a) trials in the non-ICA-corrected VEOG 25 ms before and 225 ms after word presentation with peak-to-peak thresholds exceeding 150 µV (window size = 175 ms, window step = 10 ms), and (b) trials in the non-ICA-corrected HEOG 50 ms before or 250 ms after word presentation with step-like artifacts exceeding 64 µV (window size = 100 ms, window step = 10 ms). After implementing this procedure, the overall mean rejection rate was 21.48% (*SD* = 22.57%), with a per participant average of 209.12 (*SD* = 70.48) accepted trials in Block 2 of the experiment, for which my primary analyses are based. These preprocessing steps align generally with those in Batterink et al. (2014).

3.3.5 Debriefing questionnaire

Following the main experimental task, a structured interview was administered to assess the extent of participants' rule awareness, following the questions given in Appendix D. Participants were first asked if they had noticed any pattern about when the different articles were used, beyond the overtly taught near/far rule. If at this point participants spontaneously reported that certain articles co-occurred with living/nonliving referents more often than others, participants were asked at what point they had noticed this pattern (i.e., during the first half of the experiment prior to the break, during the second half of the experiment following the break, or only when directly asked during the interview). Whether or not participants reported that they had noticed any patterns, they were then asked whether they had been looking for rules/patterns during the experiment, told that an underlying rule determined article usage in most cases, and were invited to guess this rule. Participants were then asked to indicate, using a continuous sliding scale without numeric labels for each of *ul/gi/ro/ne*, which articles co-occurred more often with living vs. non-living nouns.

Our original intent was to code participants as rule-aware/rule-unaware following the same criteria as Batterink et al. (2014), wherein participants who produce the correct pattern and report having noticed it during either the first or second experimental block were classified as rule-aware. Specifically, they state:

> Participants who were unable to accurately describe the pattern even after prompting, or who reported not becoming aware of the pattern until being directly questioned about it during the interview stage were classified as Rule-Unaware. Participants who described the pattern more or less accurately and who reported becoming aware of the pattern while performing the online experimental task (during either the Pre-Nap or Post-Nap session) were classified as Rule-Aware (p. 171, Batterink et al., 2014).

However, when following this approach, only very few of my participants (3 out of 52) met this criterion exactly. This may have been due to differences in the experiment designs (e.g., Batterink et al.'s participants had a nap separating Blocks 1 and 2 whereas this dissertation's participants only got a five-minute break, which is particularly relevant given that Batterink et al. reported that sleep facilitated learning; Batterink et al.'s participants saw the pseudoword article on the same page as the noun whereas my study presented these sequentially, making the task potentially more difficult because participants had to keep the pseudoword article in memory for longer while providing the task responses; my stimuli included words referring to humans in the category of living things, making a pattern less difficult to detect than when just including non-human animals; etc.) which made the rule harder to notice in my design. In order to allow for a systematic comparison between roughly equally sized participant groups with higher vs. lower levels of awareness, I relaxed the criteria such that participants were classified as rule-aware if they mentioned animacy as a relevant feature either when asked if they had noticed any patterns in the experiment or when asked to guess the underlying rule, regardless of whether they reported noticing it during Block 1 or Block 2. In this sense, my coding of rule-awareness essentially distinguishes "more implicit" vs. "more explicit" participants, following the intuition that gradient levels of awareness made the relevant feature more salient to the rule-aware participants during the debriefing interview.

To illustrate this more tangibly, some illustrative guesses at the hidden rule from participants who were coded as rule-aware include: "rule may be in regards to whether it's living or nonliving as in animal or object"; "gi for living and ro for nonliving like, for example, gi runner and ro plastic"; "ro could be used to refer to objects that are nonliving, ne was always used with nonliving objects, and ul was for both"; "a lot of animals had gi."; and "like *el* or *la* in Spanish, kind of like noun gender based on whether it's living or nonliving." Some illustrative rule guesses from participants who were coded as rule-unaware include: "maybe something about vowels or consonants"; "maybe after using gi and ro then you'd go back to ul or some similar pattern in terms of cycling through the articles"; "something to do with the colors from the pretraining, like gi was red and ro was blue, ul was gray and ne was green"; and "ro is for proper nouns and ul is for proper nouns, and gi and ne are for nonproper nouns." My rule-awareness coding was performed for each participant by two independent raters. Wherever disparities in coding occurred, these were resolved by discussion between the raters. This coding showed high interrater reliability with a Cohen's kappa of 0.81, which is categorized as "almost perfect" as per the standard interpretation guidelines in Landis and Koch (1977).

In our debriefing interview data, there were no cases of participants listing as relevant criteria attributes that correlated with living/nonliving categories, e.g., using *gi* and *ul* for "gross" words like "spider" and "snake," as found in previous experiments using this artificial language paradigm (Phillip Hamrick, personal communication). This may be because the living and nonliving words that we used were matched for semantic attributes like concreteness and valence.

3.4 Analysis

3.2.1 Behavioral analysis

As in Batterink et al. (2014), the principal behavioral measure of learning of the covert animacy rule was the Rule Learning Index (RLI), which comprises response time slowdowns for the animacy response in violation trials when compared to rule-adhering trials. More specifically, data from each of the three experimental blocks were divided into four epochs of equal length, yielding twelve total epochs. To examine rule learning prior to the debriefing interview, participants' median response times (calculated per epoch from raw, non-transformed scores) to rule-adhering vs. rule-violating trials in Blocks 1 and 2 were compared using a Greenhouse-

Geisser-corrected mixed 2x2x8 ANOVA, with Awareness (rule-aware vs. rule-unaware) as a between-participants factor and Condition (rule-adhering vs. rule-violating trial) and Epoch (for each of eight experimental epochs) as within-participant factors. Any significant effects involving two-level factors were followed up via Bonferroni-corrected t-tests, and any significant effects involving the eight-level factor Epoch were followed up via a linear trend analysis. Only trials with correct responses to the animacy judgment were included in this analysis. Trials with responses slower than 3 seconds were trimmed from the analysis. All analyses were performed in R using the *aov_car*, *emmeans*, and *effsize* packages, and results plotted with the *ggplot2* package.

As in Batterink et al. (2014), a secondary behavioral analysis was conducted on participant accuracies to the animacy response. This followed the same approach as the response time analysis, except that mean accuracy (calculated per participant per epoch per condition) was used as the dependent variable rather than median response time.

There are certain statistical issues that stem from the fact that we had six times as many rule-adhering trials as rule-violating trials. For instance, Miller (1988) notes that the use of sample medians across unequal sample sizes may lead to overestimation of population medians when distributions are positively skewed, as is generally the case with reaction time data. However, re-analyses of Batterink et al.'s (2014) original reaction time data showed that the learning effect was robust even when implementing a bootstrap-based bias-correction technique proposed by Rousselet and Wilcox (2020) (Abugaber & Morgan-Short, 2020). This suggests that the implicit learning effect is robust when accounting for unequal sample sizes, at least at a behavioral level.

3.2.2 RQ1: ERP analysis

For Research Question 1 (*Can event-related potentials capture differences in learning of a covert morphosyntactic regularity between rule-aware and rule-unaware learners?*), the ERP analysis reported in Batterink et al. (2014) was replicated as follows. Trial epochs extracted from -200 to 1200 milliseconds relative to the onset of the noun were averaged for each of the rule-adhering and violation conditions. As in Batterink et al. (2014), the time windows of analysis were 400-800 milliseconds (capturing the first reported ERP effect, an early negativity) and 800-1100 milliseconds (capturing the second ERP effect, a late positivity). For each of these two time windows, both a midline analysis and a lateral analysis were performed. For the midline analysis, mean amplitudes at electrodes Fz, Cz, and Pz were computed and entered into a mixed Greenhouse-Geisser ANOVA with Condition (rule-adhering vs. rule-violating trial) and Electrode (Fz, Cz, or Pz) as within-participant factors and Awareness (rule-aware vs. rule-unaware) as between-participant factors. For the lateral analysis, mean amplitudes at electrodes F7, F3, F4, F8, FC5, FC1, FC2, FC6, T7, C3, C4, T8, CP5, CP1, CP2, CP6, P7, P4, P3, and P8 were computed and entered into a mixed Greenhouse-Geisser ANOVA with Condition (2 levels: rule-adhering trial, rule-violating trial), Hemisphere (2 levels: left, right), Anteriority (3 levels: anterior, middle, posterior), and Laterality (2 levels: central, lateral) as within-participant factors and Awareness (2 levels: rule-aware vs. rule-unaware) as a between-participant factor. Because my ERP indices of interest were specifically oriented towards learning of the implicit rule (which would be reflected as differences between rule-adhering vs. rule-violating trials), only significant effects involving the Condition factor were analyzed in follow-up tests, using Bonferroni-corrected t-tests whenever an interaction involved more than three levels of any factor. Effect sizes are reported as generalized eta squared (η^2_G), which represents the amount of variance explained by a given model term while presenting a standardized measure across both within-and between-participant research designs (Olejnik & Algina, 2003). As in Batterink et al. (2014), this analysis only included correctly responded trials from the second block, after sufficient time would have passed in the experiment for learning to occur.

Although our rule-adhering vs. rule-violating trial conditions have a different number of trials, because we are measuring mean amplitude rather than peak amplitude then this is not a significant issue, as discussed in Chapter 9 of Luck (2014).

3.2.3 RQ2, RQ3: MVPA analyses

This section presents an overview of the Multivariate Pattern Analysis (MVPA) method as well as the specific analysis decisions for each of Research Question 2 (*Do rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data?*) and Research Question 3 (*Do rule-aware and rule-unaware participants show neural evidence of semantic prediction from a covert morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so do they differ in their use of prediction?*). As discussed previously in Section 2.3.5, MVPA involves training a decoder to classify trials from different experimental conditions by automatically finding the decision boundary that best distinguishes trials from these conditions when these data points are projected in a multidimensional way, i.e., with each electrode constituting a separate dimension and each data point corresponding to the simultaneous activation level in each of the electrodes for a given trial. An illustration of this is given in Figure 3, showing only two dimensions for simplicity.

As can be seen in the figure, at times a traditional ERP analysis may be unable to detect differences between conditions due to substantial overlap between the data points within any single electrode. However, by taking into account correlations with other electrodes, one could find an underlying boundary that best separate data points from the different conditions in a statistically significant way. The dimension of this newly found decision boundary is itself an abstract amalgamation of voltages from several electrodes, rather than the simple positivity/negativity of voltage from within any given single electrode (as in a traditional univariate analysis). This is the principle that underlies MVPA decoding.

After MVPA decoder training, testing of the decoder's trial classification accuracy is performed on separate data, which may be either data that was withheld from the original dataset during training or an entirely separate dataset altogether. In either case, above-chance accuracy in classifying trials from the testing data (based on where these individual data points lie in relation to the decision boundary) would thus suggest that similar neural activation patterns are involved between the training data and the test data (Kaplan, Man, & Greening, 2010). Using this approach, one can also make inferences about whether patterns of neural activity from one time point are also found at other time points, i.e., by seeing whether the decision boundary calculated for one time point also shows above-chance accuracy when applied to different time points (King & Dehaene, 2014). This temporal generalization is critical for pulling apart ERP components that occur in temporally overlapping windows (as might be the case in Batterink et al., 2014). The temporal generalization method is illustrated in the analysis pipeline shown in Figure 4.

There are several analysis decisions that can affect the performance of an MVPA classifier. For instance, given the high dimensionality of EEG data (which comprises dozens of individual electrodes multiplied by hundreds of time points per trial) as well as the correspondingly strict corrections for multiple comparisons necessary to avoid Type I error rate inflation, one can improve signal-to-noise ratio (and, consequently, classifier performance) by selecting the electrodes of analysis *a priori*, based on prior knowledge of the effects observed in prior univariate analyses (Grootswagers et al., 2017). For my analyses, all scalp electrodes were included because we had no strong predictions regarding the location of effects (though see the section titled “MVPA results when using region-of-interest electrodes” in Appendix E). Additionally, all decoder training and testing were performed at the level of individual participants rather than on aggregated data sets. This per-participant level of MVPA analysis means that each participant had their own decoder with its own decision boundaries tailored to their dataset, which can help to overcome issues arising from individual differences in language processing signatures (e.g., Tanner & Van Hell, 2014). This contrasts with traditional ERP approaches that aggregate data across participants and analyze them at a group level, under the assumption that group averages are generally representative of individuals.

All MVPA analyses were performed using the Amsterdam Decoding and Modeling Toolbox (ADAM; Fahrenfort et al., 2018), a freely available MATLAB toolbox for EEG and MEG data. For these analyses, correction for multiple comparisons was performed using a cluster-based permutation method. More in-depth details about this method as well as my other analysis parameters for MVPA are given in Appendix E in the section titled “Design decisions for results presented here.” In the subsections below, I describe the specific approach for the MVPA analyses that were performed for the extensions of Batterink et al. (2014) from Research Questions 2 and 3.

3.2.3.1 Distinguishing neural signatures of implicit vs. explicit processing

To address Research Question 2—*do rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data?*—we aimed to train a decoder to detect measures of rule-unaware processing and subsequently test this decoder on periods of rule-aware processing. For participants who were classified as rule-unaware in my debriefing interview, I train my MVPA decoder to distinguish rule-adhering vs. rule-violating trials in the second block (after learning would be more likely to occur). Then, I tested this decoder's accuracy in classifying trials as rule-adhering vs. rule violating in data from the third block (after participants heard the rule from the experimenter at the end of the rule-awareness debriefing). This allows us to determine whether neural activity during periods of rule-unaware grammar processing extends to neural activity during periods of rule-aware processing. As in the univariate ERP analyses, only trials with correct responses to the animacy judgment were used, following Batterink et al. (2014).

Note that this approach performs all analyses within-participant rather than training an MVPA decoder on one group of participants' data and testing it on another group's data. The reason I did not take this approach is that there are several potential factors that may "stack the deck" against inter-participant MVPA, not least among them individual differences in brain anatomy (Kaplan & Meyer, 2012), in reported differences in individuals' ERP signatures (e.g., Tanner & Van Hell, 2014), and in the attainment of detectable P600 responses in an L2 processing context (e.g., Bond et al., 2011). Furthermore, because this MVPA decoder would be trained on a compilation of data from all rule-unaware participants (rather than on a per-participant basis), any reliable criteria that the MVPA decoder might have used to detect implicit processing within a given individual data set—which might be participant-specific due the aforementioned individual differences in neurolinguistic processing—might "wash out" when trials are aggregated across participants to train a single decoder. In light of these obstacles, a cross-participant MVPA analysis was not performed.

3.2.3.2 Detecting neural indices of semantic prediction with MVPA

To address Research Question 3—*do rule-aware and rule-unaware participants show neural evidence of semantic prediction from a hidden morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so do they differ in their use of prediction?*—the first step was to train an MVPA decoder to classify trials for animate vs. inanimate nouns in the experiment. This would give us a measure of neural activity that is specifically oriented to animacy (the "predicted-for" semantic feature), which could subsequently be tested on neural activity evoked by the reading of the artificial language article (as described further below). I trained such a decoder on the noun-only portion of the experiment (see Section 3.1.3 above).

This animacy-trained MVPA decoder was subsequently tested on epochs extracted in the time window -200 to 1200 milliseconds relative to the presentation of the artificial language article. If the animacy-trained decoder shows above-chance accuracy in classifying these artificial language trials—that is, in determining whether the artificial language article was animacy-encoding (that is, *gi* or *ul*) or inanimacy-encoding (that is, *ro* or *ne*), based simply on the detection of neural patterns tied to the activation of animacy features—then this would indicate prediction for an (in)animate entity based simply on the reading of the artificial language article. Beyond just determining whether prediction occurs in the first place within a given participant group, I also aimed to determine whether these two groups differed in prediction by performing t-tests with cluster-based correction for multiple comparisons across the rule-aware vs. rule-unaware participants on their respective classification accuracies for each time point in the testing data. In this way, MVPA was utilized to determine whether linguistic prediction differentially characterizes rule-aware and rule-unaware processing.

3.2.4 RQ4: Reaction time-to-ERP correlations

To address Research Question 4—*is the P600 component time-locked to participants' external behavioral responses in an artificial language learning task?*—several different analyses are performed. This decision to use separate, potentially redundant analyses rather than a single analysis follows the approach taken in a previous study that specifically aimed to quantify the response-locked nature of the P600 component (Sassenhagen et al., 2014), which noted that no single standard method for quantifying the alignment of reaction times with ERP components has been established. Just as in this prior study (Sassenhagen et al., 2014), four different analyses were performed in this dissertation.

The first analysis involved the so-called ERP image technique (Jung et al., 2001), which is a visualization of component latencies rather than a formal test that uses inferential statistics. As illustrated in Figure 5, in this method event-locked EEG epochs are "stacked" horizontally as color-coded lines (with color indicating voltage), with trial time on the *x* axis and individual trials as separate bands on the *y* axis. By sorting these trials on the *y* axis based on rule-aware participants' response times, an image emerges wherein the P600 effect's temporal relationship to response time can be seen. Specifically, if the P600 component is purely aligned to the stimulus onset, then the observed color bands associated with the P600's positivity would occur at roughly the same time for all trials regardless of response time, following a roughly vertical pattern in the response time-sorted ERP image plot. Conversely, if the P600 component is temporally related to participant response, then a diagonal pattern would emerge wherein the effect occurs earlier for trials with earlier responses (i.e., the trials at the bottom of the ERP image plot), and the effect occurs later for trials with later responses (i.e., the trials at the top of the ERP image plot).

The second analysis of the relationship between response time and P600 latency involves response time binning (Poli et al., 2010; Marathe et al., 2013), which has previously been used to correlate response time with other ERP components such as the P3 (Marathe et al., 2013; Poli et al., 2010; Roth et al., 1978). In this approach, each rule-aware participant's trials are binned according to that participant's response time quartile, i.e., divided into four categories comprising the trials with the fastest, middle-fastest, middle-slowest, and slowest response times, calculated per participant. Subsequently, the latency of the ERP P600 component is estimated per bin, and the bins are then ranked by latency. If P600 latency increases with bin rank, then a relationship between the component and response time can be inferred. Following standard procedures (Kiesel et al., 2008; Luck, 2014; Ulrich & Miller, 2001), in my analysis I exclude outlier trials with the 2.5% lowest and 2.5% highest reaction times (calculated per participant); set all negative values to zero to avoid contributions from a negative component; construct jackknife averages to increase the stability of my findings; and estimate P600 component onset by using 33% fractional latency of the area under the positive curve. Repeated-measures one-way ANOVAs on P600 latencies with the factor of reaction time quartile were conducted to test for a statistically significant effect.

In my third approach to measuring the relationship between reaction time and P600 component latency, I calculate correlations between response times and individual trial component latencies following the Woody filtering procedure previously used by Kutas et al. (1977) and Sassenhagen et al. (2014). First, each rule-aware participant's mean difference ERPs for rule-conforming vs. rule-nonconforming trials was calculated. Then, the time lag of the best correlation between this per-participant ERP template and each of that participant's trials was measured. For 100 iterations, a new template ERP was calculated by shifting individual trials by the lag that I had identified, and subsequently I computed the best correlation between the ERP template and individual trials. The latency of the P600 effect was calculated as the point of best correlation between single trials and the final template iteration. Finally, Pearson's correlation coefficients between single-trial response times and this P600 latency were calculated for individual participants, such that high *r* values would indicate a strong correlation (and therefore an underlying relationship) between response times and the P600 component.

Our fourth and final analysis of the relationship between response times and the P600 component involves calculating inter-trial phase coherence (ITC; Delorme et al., 2007), which provides a measure of the cross-trial phase consistency of EEG oscillations. This was done by Fourier-transforming each trial (i.e., decomposing the observed waveform into its underlying wavelets) and computing the consistency of the oscillation phase per frequency and per time point across trials. By comparing each trial's data to two different temporal alignments—one corresponding to stimulus onset and another corresponding to participant response time, I aimed to determine with which of these two time points the P600 component is better aligned. Following previously validated procedures (Sassenhagen et al., 2014), I calculated the time and frequency mean ITC from 0.5 to 8Hz in a 50 ms time window centered on the positive peak of the P600 for each participant.

In this way, for Research Question 4 this dissertation study performed four separate analyses for determining the relationship between the P600 ERP component and behavioral response times, following the approach taken in a previous study that had precisely the same goal (Sassenhagen et al., 2014). In the interpretation of my results, no single analysis was be prioritized over another, just as no single standard method for reaction time-based analyses has emerged in prior ERP research. Rather, the results from these four analyses were considered together to provide a complementary picture for addressing Research Question 4. Any disparities between the results of these analyses were examined and are discussed below in light of the particular statistical characteristics and parameters involved.

CHAPTER IV. RESULTS

Of the 52 recruited participants, four were excluded. Two were excluded due to technical issues with the experiment; one was excluded because they decided to quit the experiment early due to excessive fatigue; and one was removed for low accuracy in the behavioral task (mean accuracy for living/nonliving response = 53.28%). Following removal of these participants, mean accuracy for the living/nonliving response in my task was 87.92% (*SD* = 6.75%, range = 71.23-96.39%). This suggested that participants were attentive and performing the task successfully. Of the 48 participants included in my final dataset, 24 were coded as (more) rule-aware and 24 were coded as (less) rule-aware.

4.1 Behavioral results

4.1.1 Response time analyses

A response time analysis was conducted to test for a learning effect using the same analysis procedure as in Batterink et al. (2014), focusing on Blocks 1 and 2 (prior to the rule debriefing questionnaire). Figure 6 shows epoch median response times to correct (i.e., rule-adhering) vs. violation (i.e., rule-violating) trials, calculated per participant. Descriptively, response times seemed to become faster over the course of the first block before leveling off at the start of the second block and remaining low for the rest of the experiment. Furthermore, violation trials seemed to show slower response times compared to correct trials for each epoch starting halfway through the first block. In order to test this formally, a Greenhouse-Geisser-corrected mixed 2x2x8 ANOVA was performed on median reaction times (calculated per participant per epoch) with Awareness (rule-aware vs. rule-unaware) as a between-participant factor and Condition (rule-adhering vs. violation) and Epoch (for each of the eight experimental epochs in Blocks 1 and 2) as within-participant factors. As shown in Table III, this yielded a significant main effect of Condition, $F(1,46) = 15.17$, $p < .001$, $\eta^2_G = .01$, and of Epoch, $F(3.67, 168.89) = 54.99$, $p < .001$, $\eta^2_G = .24$, but not of Awareness, $F(1,46) = 0.03$, $p = .864$, $\eta^2_G < .01$. There were no significant interactions with Awareness (all ps > .05). Thus, no evidence was found for a difference in the learning effect between rule-aware and rule-unaware participants. Although there were no main effects or interactions from Awareness, for exploratory purposes I visualize the reaction time results for rule-aware and rule-unaware participants in Figures 7 and 8, respectively.

A follow-up t-test on the main effect of Condition in Blocks 1 and 2 found that epoch median response times to rule-violating trials (M = 1113.96 ms, SD = 359.05) were significantly slower than to rule-adhering trials (M = 1063.55 ms, SD = 300.58), $t(46) = 3.90$, $p < .001$, d = 3.91. A follow-up trend analysis performed for the main effect of the multilevel factor Epoch found a significant linear trend, $t(322) = -16.62$, $p < .001$, $\beta = -4.97$, such that later epochs had significantly faster reaction times than earlier epochs, averaged over Condition or Awareness.

For exploratory purposes, I report here a similar analysis of reaction times in Block 3 (after participants had been explicitly told the hidden rule). I performed a Greenhouse-Geisser-corrected repeated-measures 2x2x4 ANOVA of participants' epoch median reaction times with Condition (correct vs. violation), Rule Awareness (rule-aware vs. rule-unaware), and Epoch (for each of four experimental epochs in Block 3) as within-participant factors. As shown in Table IV, this yielded a significant main effect of Condition, $F(1,46) = 38.60$, $p < .001$, $\eta^2_G = .08$. There were no significant main effects or interactions with Awareness (all ps > .05), As such, I found no evidence for a significant difference in the learning effect between rule-aware and rule-unaware participants in Block 3. Follow-up t-tests on the main effect of Condition found that epoch median response times to rule-violating trials (M = 1036.97 ms, SD = 341.58) were significantly slower than to rule-adhering trials (*M* = 868.10 ms, *SD* = 246.30), $t(46) = 6.213$, $p < .001$, $d = 6.19$.

In sum, our reaction time analysis revealed a rule-learning effect such that rule-violating trials showed slower responses than rule-adhering trials, with no effects or interactions from rule-awareness.

4.1.2 Accuracy analyses

To supplement the reaction time analysis, an accuracy analysis was conducted to test for a learning effect using the same analysis procedure as in Batterink et al. (2014) for Blocks 1 and 2 (prior to the rule debriefing questionnaire). Figure 9 shows participants' mean epoch accuracies to rule-adhering vs. rule-violating trials. Descriptively, accuracies seemed to be near ceiling throughout the experiment, with the exception that violation trials showed a slight reduction in the later epochs. To test this formally, a Greenhouse-Geisser-corrected mixed 2x2x8 ANOVA was performed on participants' mean epoch accuracies with Awareness (rule-aware vs. rule-unaware) as a between-participant factor and Condition (rule-adhering vs. rule-violating trial) and Epoch (for each of the eight experimental epochs in Blocks 1 and 2) as within-participant factors. As can be seen in Table V, this yielded a significant main effect for Condition, $F(1,46) = 5.34$, $p = .025$, $\eta^2_G = .01$, but no other significant effects. Importantly, there were no significant interactions with Awareness (all *p*s > .05), meaning that no evidence was found for a significant difference in the learning effect between rule-aware and rule-unaware participants. Although there were no main effects or interactions from Awareness, for exploratory purposes I visualize the accuracy results for rule-aware and rule-unaware participants in Figures 10 and 11, respectively.

Follow-up t-tests on the effect of Condition found that participants' epoch mean accuracies to rule-violating trials (*M* = 87.49%, *SD* = 12.55%) were significantly lower than to rule-adhering trials (*M* = 89.00%, *SD* = 7.99%), $t(46) = -2.31$, $p = .025$, $d = 2.29$.

As before, I report an exploratory analysis of Block 3 (after participants had been told the rule in the rule-debriefing questionnaire). Using a similar procedure as above, a Greenhouse-Geisser-corrected mixed 2x2x4 ANOVA was performed on participants' mean epoch accuracies with Awareness (rule-aware vs. rule-unaware) as a between-participant factor and Condition (correct vs. violation) and Epoch (for each of the four experimental epochs in Block 3) as within-participant factors. As can be seen in Table VI, this yielded a significant main effect for Condition, $F(1,46) = 37.97$, $p < .001$, $\eta^2_G = .19$, but no other significant effects. Follow-up t-tests on the effect of Condition found that participants' epoch mean accuracies to rule-violating trials (*M* = 72.29.%, *SD* = 23.97%) were significantly lower than to rule-adhering trials (*M* = 89.45%, *SD* = 8.25%), $t(46) = -6.16$, $p < .001$, $d = 6.19$.

In sum, our accuracy analysis revealed a rule-learning effect such that rule-violating trials showed less accurate responses than rule-adhering trials, with no effects or interactions from rule-awareness.

4.2 RQ1 ERP results

Grand-averaged event related potentials comparing responses to rule-adhering vs. rule-violating trials are shown for Block 1, Block, and Block 3 in Figure 12 below, both overall as well as separately for rule-aware and rule-unaware participants. Figure 12 also shows scalp maps for the time windows of interest based on Batterink et al. (2014). As can be seen, the only prominent effect seems to be a positivity for rule-unaware participants in Block 2.

To formally test these observations via inferential statistics, I turn to my analyses of variance. For each of Blocks 1, 2, and 3, I ran a midline 2x2x3 Greenhouse-Geisser-corrected ANOVA with the between-participant factor Awareness (2 levels: rule-aware vs. rule-unaware) and the within-participant factors Condition (2 levels: rule-adhering vs. rule-violating trial) and Electrode (3 levels: Fz, Cz, and Pz). I did not find any statistically significant effects for Block 1 and Block 3. For Block 2, my analysis did not reveal any statistically significant effects or interactions for the 400-800 ms time window (*p*s > .05) involving the critical factor Condition. For the 800-1100 ms time window in Block 2, the only statistically significant finding involving Condition was the interaction of Awareness by Condition, $F(1, 46) = 4.07$, $p = .050$, $\eta^2_G = .08$. Follow-up t-tests revealed that, for rule-unaware participants, amplitudes to violation trials (M = -0.19 mV, SD = 7.13) were significantly more positive than to rule-adhering trials (M = -1.74 mV, SD = 5.30), $t(46) = 1.55$, $p = .021$, $d = 2.40$, confirming the positive effect seen in the waveforms (see Figure 12). However, for rule-aware participants, amplitudes to violation trials (M = -2.17 mV, SD = 7.38), were not significantly different from rule-adhering trials (M = -1.88 mV, SD = 5.91), $t(46) = -0.45$, $p = .653$, $d = 0.45$. The ANOVA results are presented in Table VII.

Secondly, a Greenhouse-Geisser-corrected ANOVA was performed on the lateral electrodes with the between-participant factor Awareness (2 levels: rule-aware and rule-unaware) and the within-participant factors Condition (2 levels: rule-adhering vs. rule-violating trial), Anteriority (3 levels: anterior, middle, and posterior); Hemisphere (2 levels: left, right) and Laterality (2 levels: central, lateral). For Block 1, I found a significant interaction of Awareness by Condition by Hemisphere in the 400-800 ms time window, $F(1,46) = 4.06$, $p = .050$, η^2_G = .08. Follow-up analyses revealed that for rule-aware participants, amplitudes in the right hemisphere were more positive for violation (*M* = -1.63 mV, *SD* = 4.44) than for rule-adhering (*M* = -2.32 mV, *SD* = 4.36) trials, but this effect was not significant in a follow-up t-test, $t(50.4) = 1.34$, $p = .187$, $d = 1.34$. All other follow-up comparisons per hemisphere per trial condition per participant awareness status were non-significant (*p*s > .837). For Block 2, I found a significant interaction of Condition by Hemisphere in the 800-1100 ms time window, $F(1,46) = 4.39$, $p = .042$, $\eta^2_G = .09$. Follow-up analyses revealed that amplitudes in the right hemisphere were more positive for violation (*M* = -0.42 mV, *SD* = 5.39) than for rule-adhering (*M* = -1.25 mV, *SD* = 4.22) trials, but this effect was not significant in a follow-up t-test, $t(54.1) = 1.96$, $p = .056$, $d = 1.95$. Rule-violating and rule-adhering trials were not significantly different in the left hemisphere (*p*s = .451). For Block 3, I found a significant interaction of Awareness by Condition by Hemisphere in the 800-1100 ms time window, $F(1,46) = 4.22$, $p = .046$, $\eta^2_G = .08$. Follow-up analyses revealed that for rule-aware participants, rule-violating trials were more negative than rule-adhering trials for both the left and right hemispheres, but not significantly so (*p*s > 0.22). Meanwhile, for rule-unaware participants, violation trials were more negative than rule-adhering trials in the left hemisphere but more positive than rule-adhering trials in the right hemisphere, though not significantly so (*p*s > .446). No other significant effects involving the factor Condition were found in the lateral ANOVA. The full lateral ANOVA results are shown in Table VIII.

In sum, the only statistically significant effect that we found was a late positive deflection in the 800-1100 ms time window in response to rule-violating trials within Block 2 in the rule-unaware group. By contrast, no significant ERP effects were found for the 500-800 ms time window, for Blocks 1 and 3, or for rule-aware participants in any time window in any block.

4.3 RQ2 MVPA results on occurrence of implicit processing during rule awareness

Recall that my second research question aimed to determine whether rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data. To do so, I trained an MVPA decoder to distinguish rule-adhering vs. rule-violating trials in rule-unaware participants' Block 2 data (after enough time would have passed for the rule to have been learned, as evidenced by significant effects of rule-adhering vs. rule-violating trial condition in both the reaction time ANOVA and the accuracy ANOVA; see Tables III and V and Figures 8 and 11), and then tested the accuracy of categorizing rule-adhering/rule-violating trials on those same participants' data in Block 3, after the rule was revealed to them during the debriefing interview. Panel A of Figure 13 below shows results of the implicit decoder in performing diagonal decoding—that is, when each time point in the training data is tested against its corresponding time point in the testing data (for instance, the data from 300 ms is used to train the decoder and is subsequently tested at time point 300 ms; then the data at 301 ms is used for training and tested against time point 301 ms; etc.). This provides the most basic measure of the decoder's performance, without any generalization across time windows. As can be seen, slightly above-chance trial classification accuracy occurred in intermittent time periods starting at roughly 300 ms. However, the results were not statistically significant after cluster-based corrections for multiple comparisons was performed. Panel B of Figure 13 shows decoding performance on each millisecond of the testing data, when training data is limited to the time window 800-1100 ms (the time window of the late positive component detected for these participants in my midline ANOVA above and by Batterink et al., 2014). No time point showed statistically significantly above-chance decoding performance. Panel C of Figure 13 shows decoding performance in a temporal generalization matrix. This provides a visualization of how each time point in the data generalizes to every other time point in the trial, as per the Generalization Across Time approach (King & Dehaene, 2014). For example, the data from time point 300 ms would be used to train a decoder, which would subsequently be tested on each individual time point in the epoch (from 1 ms to 1000 ms); then the same would be done for time point 301 ms, time point 302 ms, etc. This allows for a broader analysis than the diagonal decoding results because decoder testing is not limited to specific time points but rather all time points in the epoch, such that if neural activity from, e.g., 300 ms also occurs at 350 ms, then this would be captured in the decoder's results. As can be seen in Panel C of Figure 13, slightly above-chance accuracies (as indicated by redder splotches) fell into a relatively square pattern starting at around 500 ms, which suggests a single, sustained process throughout the trial rather than a series of brief cascading processes (which would instead be represented by above-chance decoding accuracy along the diagonal line; King & Dehaene, 2014). However, these were not significant after cluster correction for multiple comparisons, suggesting no co-occurrence of neural activity related to the hidden rule across periods of rule-awareness and rule-unawareness.

In sum, our MVPA analysis did not yield evidence for the co-occurrence of implicit processing during rule awareness.

4.4 RQ3 MVPA results on semantic prediction in implicit vs. explicit processing

Recall that my third research question aimed to determine whether rule-aware and rule-unaware participants show neural evidence of semantic prediction from a covert morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so whether they differ in their use of prediction. To do this, I trained and MVPA decoder on a "noun-only" period of the experiment during which participants saw nouns in isolation and had to judge them as either animate or inanimate, with button responses withheld until a prompt appeared (with jittered timing so that the exact appearance was unpredictable) with button assignments randomly determined on each trial; this ensured that my MVPA decoder would be sensitive to the animate/inanimate distinction rather than to anticipation of a predictable button press. Then, this decoder was trained on the pseudoword articles from the main experimental trials in Block 2, such that above-chance accuracies would indicate that the concepts animate/inanimate were being pre-activated when the articles were presented. As can be seen in both the diagonal decoding performance (Figure 14, panel A) and the temporal generalization matrix (Figure 14, panel B), for rule-aware participants there were periods of slightly above-chance decoding performance at around 250 ms and again for short intervals between 600-1000 ms. However, these results were not statistically significant after cluster-based corrections for multiple comparisons was performed. By contrast, for rule-unaware participants there are no apparent time periods with above-chance decoding evidenced in either the diagonal decoding performance or in the temporal generalization matrix. Note that the differences in decoding performance between rule-aware and rule-unaware participants were not significant, as indicated by cluster-based permutation-corrected analyses of the difference in performance between the two groups. Thus, because decoding accuracies were not significantly above chance either for the rule-aware or rule-unaware group, we found no activation of animacy in response to the pseudowords in isolation.

In sum, our MVPA analysis did not yield evidence for semantic prediction, operationalized here as occurrence of neural activity associated with processing living/nonliving status upon reading the pseudoword in isolation.

4.5 RQ4 Reaction time-to-ERP correlation results

Recall that my fourth research question aimed to determine whether the P600 ERP component of explicit processing was time-locked to rule-aware participants' external behavioral responses in an artificial language learning task. This would allow us to determine whether the P600 (which was tied to explicit grammar processing in Batterink et al., 2014) was closely related to the production of external behavioral responses, or whether the appearance of the P600 was an epiphenomenon that was well downstream of task performance. Because the only significant ERP component that I found was a late positivity in rule-unaware participants' Block 2 data, this forms the basis of my analyses for Research Question 4.

My first approach to performing this analysis was by visualizing the ERPs using the ERP image approach, wherein the EEG amplitudes in each trial are depicted as colored bands that are stacked atop each other and sorted according to reaction time, such that a colored "stripe" of positivities parallel to the line indicating reaction times would suggest time-locking of the P600 to reaction times. Figure 15 shows ERP images for electrode FC2, for which the positive component was strongest as seen in the scalp maps in Figure 12. These ERP images suggest that, for many trials, positive deflections follow reaction times in a time-locked fashion. This can be seen in the parallelism between the reaction time line and the red diagonal stripe representing positive EEG amplitude deflections in Figure 15, panel A. Time-locking is also evidenced by the fact that variability in the timing of these positive EEG amplitude deflections disappears when the data are plotted in a "response-locked" manner, as seen by the fact that the red splotches indicating positive deflections occur with the same timing (i.e., in a vertical column rather than a diagonal stripe) when the data in the ERP image are shifted to account for their reaction times (see Figure 15, panel B). However, we note that visual inspection of ERP images is somewhat subjective, and turn now to the results to our inferential statistical tests.

Our second approach to determining whether reaction times were correlated to the timing of the late positive component involved response time binning (Poli et al., 2010; Marathe et al., 2013), wherein each participant's trials are binned into quartiles according to that participant's response times and then correlated with the latency of the ERP component per bin. I estimated ERP component onset by using 33% fractional latency of the area under the positive curve centered on the 800-1100 ms time window where the effect was found. A one-way ANOVA on ERP latencies in each of the four response time bins revealed a significant effect of bin rank, $F(3, 48) = 21.81$, $p < .001$, $\eta^2_G = .53$. Subsequent inspection of the average component timing per bin showed that ERPs generally occurred later as response times became slower (mean latency in quartile with fastest reaction time = 895.82 ms, *SD* = 4.85; mean latency in quartile with second-fastest reaction time = 941.08 ms, *SD* = 48.31; mean latency in quartile with second-slowest reaction time = 938.33 ms, *SD* = 36.00; mean latency in quartile with slowest reaction time = 1022.20 ms, *SD* = 61.56). These results suggest time-locking between reaction times and the late positive component.

Our third approach to measuring the relationship between reaction time and ERP component latency involved calculating correlations between response times and individual trial component latencies following a Woody filtering procedure (as in Kutas et al., 1977; Sassenhagen et al., 2014). After calculating a Pearson's correlation coefficient between single-trial response times and component latency for each rule-unaware participant in the Block 2 data, the mean *r* value was 0.10 (*SD* = 0.26, 95% CI = -.03 to 0.23). A two-tailed one-sample t-test found that this value was not significantly different from zero, $t(17) = 1.63$, $p = .121$. As such, this analysis does not yield significant evidence that reaction times and component latencies were calculated. Note that six of participants were dropped in this analysis because the toolbox used (Pernet et al., 2013; used in Sassenhagen et al., 2014) required a higher number of trials per participant, leaving only 18 of the 24 rule-unaware participants' data for this analysis.

Our fourth and final analysis of the relationship between response times and the P600 component involved calculating inter-trial phase coherence (ITC; Delorme et al., 2007), which provides a measure of the cross-trial phase consistence of EEG oscillations. I found that, for rule-unaware participants' Block 2 data, inter-trial coherence was slightly greater when trials were reaction time-locked than when they were not reaction time-locked (mean difference = 3.24%, *SD* = 8.66%). However, this was not found to be significantly different from zero as per a two-tailed one-sample t-test, $t(23) = 1.83$, $p = .080$.

In sum, our analysis of reaction time-to-ERP latency correlations found mixed results, with visual inspection of ERP images and a reaction time quartiling approach yielding evidence for time locking between reaction times and ERPs but Woody filtering and inter-trial phase coherence analyses yielding no statistically significant evidence for this.

CHAPTER 5. DISCUSSION

We performed a reproduction and extension of Batterink et al.'s (2014) artificial language experiment to investigate the nature of implicit vs. explicit second language grammar learning. In my reproduction component, I aimed to determine whether ERP responses can capture differences in rule-aware and rule-unaware grammar learning (Research Question 1). In my extension component, I aimed to determine whether rule-aware participants show implicit processing of the rule at a neural level (Research Question 2), whether rule-aware and rule-unaware grammar processing differ in terms of semantic prediction (Research Question 3), and whether neural signatures of explicit processing are temporally related to (and thus closely linked with) externally observable reaction times (Research Question 4). This chapter presents a general discussion of my results for each of these research questions (Section 5.1). before moving on to a discussion of the extent to which the target of learning in our study was linguistic vs. domain-general (Section 5.2). Then, the chapter goes on to discuss the relevance of our findings in light of prior literature in the fields of second language acquisition and psychology (Section 5.3) before presenting the study's limitations (Section 5.4) and a conclusion (Section 5.5).

5.1 Overall results

In terms of overall participant performance on the experiment, my behavioral results indicate that my experiment design yielded a rule-learning effect, in that reaction times were slower and accuracies were lower to trials that violated the underlying grammar rule relative to trials that followed the grammar rule. Under my approach wherein participants were categorized as "rule-aware" if they correctly guessed the nature of the rule when told that there was a hidden rule in a mid-experiment debriefing session, I attained a breakdown of 24 rule-aware and 24 rule-unaware participants in my final dataset, which is comparable to the proportions reported in prior studies using this paradigm (e.g., Batterink et al., 2014; Leung & Williams, 2011; note, however, that our criteria for rule-awareness are somewhat more relaxed in that we only required a correct rule guess rather than an explicit report of having consciously noted the rule prior to the debriefing). The sections below discuss my interpretation of the results in light of each of Research Questions 1-4.

5.1.1 RQ1 discussion

Turning to my first research question (*Can event-related potentials capture differences in learning of a covert morphosyntactic regularity between rule-aware and rule-unaware learners?*), my findings did not reproduce the results from Batterink et al.'s (2014) ERP-based analysis. Whereas Batterink and colleagues reported that grammar processing effects were manifested as a negative deflection for rule-unaware participants and a positive deflection for rule-aware participants, the only statistically significant ERP effect that I found was a positive deflection for rule-unaware participants, and this only occurred in Block 2 (before they were told the hidden rule).

To explain the disparity between my ERP results and Batterink et al.'s (2014) ERP results, I turn to the idea that natural variability between participants in ERPs may make a grand-averaged analysis inappropriate in the first place: in a study with high statistical power ($N > 100$) on native language processing in a highly homogenous participant group, Tanner (2019) reports that native speakers do not consistently show either a (negative) N400 effect or a (positive) P600 effect in response to grammatical violations, but rather fall along an N400-P600 continuum. Tanner (2019) reports that where participants fell on this spectrum was not an artifact of individual differences in language experience or verbal working memory capacity, but rather natural variation. Given this high natural variability in individual speakers, Tanner (2019) argues that aggregating data across participants and making inferences based on the resulting observed averages may not be valid because participants vary greatly between each other in the ERP effects that they show, leading to situations where, e.g., a negative-dominant participant may effectively "cancel" out the positivity from a positive-dominant participant. As such, Tanner writes that "descriptions of processing dynamics predicated solely on grand mean analyses of central tendency can fail to provide an accurate, generalizable account of how processing unfolds in many or most individual members of the population studied" (2019, p. 210). Warning about the validity of generalizing findings based on ERPs averaged across participants, Tanner et al. write that "failure to recognize important and systematic individual differences has in some cases led to inappropriate interpretations of ERP effects, with neurocognitive models of language comprehension sometimes being built on these inappropriate interpretations" (2018, p. 299).

Contextualizing this in my study, I surmise that whether a group showed a negativity or a positivity may be the artifact of which group happened to have more of one kind of ERP response along a spectrum analogous to the N400-P600 spectrum found by Tanner (2019). There are several reasons why this variability in ERPs might have been particularly pronounced for my study. As noted by Brouwer and Crocker (2017), having two co-occurring components like an N400 and a P600 within the same trial means that idiosyncratic features of the experimental task/stimuli can have different modulating effects on each of the components. For instance, different amounts of priming for the different nouns (e.g., influenced by the running list of nouns that each participant sees in the experiment) mean that the strength of the N400 response to word reading would vary for each word. In turn, differing amounts of familiarity with the task and with the response predictability from the artificial language articles could affect the timing of processes related to a late P600-like component that are initiated in the 350 ms time window after the pseudoword is shown but before the noun is shown. Seen in this way, the fact that my experiment involved a pseudoword shown before one of many possible English nouns means that many sources of variability are introduced, beyond a design where the words are shown simultaneously (as in Batterink et al., 2014) or where potentially only a single ERP component is elicited. In support of this idea that individual participants' ERPs differed greatly even within rule-aware and rule-unaware groups, I illustrate in Figure 16 how my participants fell along a natural spectrum of ERP variability just like Tanner's (2019), with no seeming systematicity based on rule awareness.

We add that our reproduction of Batterink et al. (2014) goes towards addressing a growing call for more replication studies in EEG research (Keil et al., 2014; Luck & Gaspelin, 2017) as well as more broadly in the fields of psychology (Lamal, 1990; Francis, 2012) and second language acquisition (Marsden et al., 2018).

5.1.2 RQ2 discussion

Turning to my second research question (*Do rule-aware participants show implicit processing of the rule at a neural level, as revealed by multivariate pattern analyses of EEG data?*), my analysis involved training an MVPA decoder on time periods of the experiment when participants had no apparent self-reported conscious awareness of the underlying grammar rule but nevertheless showed behavioral signs of rule learning, and subsequently testing this decoder on a final block of the experiment after the participants had been told about the hidden rule. Because my decoder did not show significantly above-chance trial classification accuracy, I did not find evidence for co-occurrence of implicit grammar processing during periods of rule awareness.

A major caveat to interpreting this “absence of evidence” as “evidence of absence” is that follow-up analyses revealed that my MVPA decoder did not show above-chance trial decoding accuracy *even when trained and tested on the same block of trials*, i.e., when the decoder was trained on rule-unaware participants’ Block 2 data and tested on that same data using five-fold cross validation (see full results presented in Appendix E, section “MVPA results in within-block decoding”). In other words, before I could hope to detect neural indices of implicit processing at the same time that explicit processing occurs, I would hope to at least be able to detect neural indices of implicit processing at times when processing is purportedly implicit. However, because my MVPA decoder was unable to detect such neural signatures, it is doubtful that my decoder would be sensitive enough for the purposes of my research question. Appendix E presents additional follow-up MVPA analyses indicating how these null decoding results reported here remain null even when using alternative analysis parameters (see sections titled “MVPA results with within-class balancing of ul/gi/ro/ne trials”, “MVPA separately on each of ul/gi/ro/ne”, “MVPA results when using region-of-interest electrodes”, and “MVPA decoding on left vs. right button presses and for each button press assignment”), and how my MVPA analysis parameters are sufficient to detect neural signs of grammar processing in native English reading data from another experiment (see section “MVPA on natural language grammar processing shows robust effects”). Given these limitations in my decoder’s trial classification performance in the current data, my MVPA findings may have limited relevance for the interface debate as to whether explicit processes help (DeKeyser, 2007), hinder (Ellis & Sagarra, 2010), or have no effect on (Paradis, 2009) the successful acquisition of implicit processing routines.

5.1.3 RQ3 discussion

Turning to my third research question (*Do rule-aware and rule-unaware participants show neural evidence of semantic prediction from a covert morphosyntactic regularity, as revealed by multivariate pattern analyses of EEG data, and if so do they differ in their use of prediction?*), my results did not find that an MVPA decoder trained on the animate/inanimate distinction could classify animacy-denoting pseudowords at above-chance accuracy, either for rule-aware or rule-unaware participant groups.

As with Research Question 2, however, one caveat here is that the relevant MVPA decoder could not accurately distinguish neural responses to animate vs. inanimate nouns when training the decoder on nouns presented in isolation in our "noun-only" block and subsequently testing the decoder on these same data using five-fold cross validation (see Appendix E, "MVPA results in within-block decoding"). In other words, because my decoder could not distinguish neural signs of processing animacy when reading isolated nouns in the first place, it is doubtful that it could detect such neural signs as a response to reading a pseudoword article. One reason for the difficulty in attaining above-chance decoding for my animacy decoder may have been that the noun-only portion of the experiment on which this decoder was trained was carefully designed to avoid "contamination" from lower-level motor processes or button press anticipation on behalf of the participant. Namely, after reading each noun, participants were shown a blank screen of randomly varying duration before the response cue indicating which button corresponded to a "living" response and which button corresponded to a "nonliving" response. This design was meant to ensure that the only difference between neural responses to living and nonliving nouns were in the living/nonliving distinction itself rather than in any lower-level factors (e.g., a left button press for "living" and a right button press for "nonliving"). Ultimately, null results in decoding the living/nonliving distinction would be preferable to spurious findings caused by lower-level confounds in trial design. Appendix E presents relevant findings suggesting that these null results were not due to parameters in the analysis but rather because the MVPA decoder simply was not sensitive enough.

5.1.4 RQ4 discussion

Turning to my fourth research question (*Is the P600 component of explicit processing time-locked to rule-aware participants' external behavioral responses in an artificial language learning task?*), my analyses yielded somewhat conflicting results. To begin with, the only significant ERP we found—and thus the only ERP for which this analysis was performed—was in rule-unaware participants' Block 2 data (the only trials in my experiment that showed a significant ERP effect). As such, my results cannot be said to reflect ERP components of *explicit* processing but rather of rule-unaware processing.

For these data, both the ERP image approach and the response time quartiling approach showed evidence for time-locking between the ERP component latency and the participants' behavioral reaction times. For the ERP images, both the stimulus-locked (Figure 15, Panel A) and the response-locked (Figure 15, Panel B) plots show red splotches (representing positive deflections) that roughly form a stripe that parallels the line indicating reaction times. Furthermore, my correlation analyses showed that the quartile of trials with the fastest reaction times tended to show earlier ERP components and vice versa. However, my other two analyses—although they trended in a direction suggestive of time-locking—were ultimately not statistically significant (for Woody filtering, confidence interval for average correlation between single-trial response times and component latency: -.03 to 0.23; for inter-trial coherence, response-locked ERPs only show improved phase coherence relative to stimulus-locked ERPs at p = .080). To the extent that these latter two analyses are more fine-grained (vs. the somewhat subjective business of "eyeballing" ERP image plots, or the rather broad approach of comparing quartiles of data rather than individual trials), I cannot say that I would expect these to be less sensitive analyses, all things being equal.

In reconciling these mixed findings, I note that the ERP components under analysis were rather faint to begin with: on the one hand, the ERP components were only barely at the threshold of statistical significance (Awareness x Condition at $p = .050$ in the Block 2 ANOVA where this effect was detected). As such, a stronger relationship between reaction times and ERP latencies might be detected if my ERP effects were more prominent in the first place. Ultimately, to the extent that the ERP components are not consistently elicited in this experiment design (e.g., due to cross-participant variability in neural responses to grammatical violations as in Tanner, 2019), then it may be premature to draw inferences from our data about these ERPs' relationship to behavioral performance. This may be in part due to the fact that a semi-artificial language learning paradigm could be expected to yield a much lower signal-to-noise ratio relative to more simplified, domain-general ERP paradigms using non-linguistic stimuli or stimuli targeting lower-level linguistic features with less variability (e.g., mismatch negativities to unexpected tones or to individual phonemes in an "oddball" paradigm, as in Näätanen et al., 1997).

5.2 Linguistic vs. low-level domain general learning in my study

To give a broad overview of my study's findings, our behavioral analysis yielded evidence of grammar learning in all participants, regardless of whether or not they had conscious awareness of the hidden rule. Our EEG findings were less pronounced: although ERP responses did indeed differ between rule-aware and rule-unaware learners (addressing Research Question 1), rule-aware learners did not show a significant ERP at all, and the significant ERP response in rule-unaware learners disappeared in Block 3 subsequent to the debriefing questionnaire. Meanwhile, my decoding-based analyses did not find evidence either for the co-occurrence of implicit processing during periods of rule awareness or for semantic prediction from the artificial language article; however, low sensitivity in my decoding results may mean that these may not be true negatives. My reaction time-to-ERP correlations provide mixed evidence for a relationship between ERP components and behavioral performance, though this may be tempered by the weak ERP effect that we detected in the first place.

My findings suggest the question: if my behavioral effect was so strong, why did I find a lack of consistent neural effects? One possible explanation comes from the nature of my trial design, which may allow ostensive linguistic processing effects to be driven by lower-level learning. Because a unique keyboard key was assigned to each of the “living” and “nonliving” responses (as in Batterink et al., 2014) and because in my experiment the pseudowords preceded the English noun, then participants may learn to anticipate the correct button press from the pseudoword without having to even read the subsequent English nouns (at least for the 87% of trials that are rule-adhering). This is because, by definition, six out of every seven trials with the pseudowords *gi* or *ul* would have a correct response on the left button (the key uniquely assigned to mean “living”). Meanwhile, six out of every seven trials with the pseudowords *ro* or *ne* would have a correct response on the right button (the key uniquely assigned to mean “nonliving”). As such, rather than *grammatical* knowledge—i.e., knowledge that *gi* and *ul* mean “living” and that *ro* and *ne* mean “nonliving”—the learning effect may instead be underlyingly manifested as knowledge that *gi* and *ul* mean “left button press” and *ro* and *ne* mean “right button press”—knowledge that could hardly be expected to manifest via typical neural signatures of language processing. In my personal experience piloting the experiment on myself as if I were a participant, I have found that by exploiting this regularity it is possible to achieve accuracies around 87% and reaction times as low as 200 ms without even reading the nouns.

This issue with fixed button assignments may also have been an issue for Batterink et al. in the first place, because participants could learn to associate *gi* and *ul* with a left button press rather than with "living", and to associate *ro* and *ne* with a right button press rather than with "nonliving" in their design as well. However, because they showed the two words on the same screen, then the predictive nature of the pseudoword might have been less stark in their experiment. Batterink et al. write that they "presumed that due to the automatic nature of reading, both the article and noun should be processed concurrently, prior to the animacy response" (2014, p. 171). By contrast, in my experiment the motor response contingency may have been more prominent because the staggered nature of pseudoword and noun presentation would give participants an opportunity to anticipate a button press ahead of time, during the 350 ms duration when the pseudoword was shown but before the noun appeared.

I present below a series of follow-up analyses that suggest a low-level domain-general learning effect in my experiment, followed by a series of arguments about why linguistic processing may indeed have been involved at least to some extent.

5.2.1 Participant debriefing responses allude to patterns in finger movements

One data point that suggests that a lower-level motor mechanism may have been involved in learning is that, in their linguistic debriefing responses, many participants reported noticing a recurring button press pattern with their fingers. For instance, some participants responded that they "developed a pattern with fingers such that *gi* was mostly the left keys"; that they "[thought] it was something subconscious like [their] fingers hovered over the next button response"; or that they recognized "memory pattern in [their] fingers" or "patterns with the button presses such that fingers moved diagonally across responses." Of our 48 participants, 10 made some reference to recurring finger patterns or to the physical layout of the correct button presses on the response keyboard during their debriefing interview (of these 10 participants, 5 were ultimately coded as rule-aware and 5 coded as rule-unaware). This issue calls into question whether the task used in this experiment truly tapped into purely linguistic processing versus simple low-level associations of words with specific button presses, in the style of a nonlinguistic serial reaction time task (e.g., Nissen & Bullemer, 1987).

5.2.2 Drift-diffusion modeling suggests motor anticipation effects

Another data point that suggests that lower-level prediction rather than higher-level processing was involved in my task comes from re-analyses of previous pilot data using drift-diffusion modeling, which can quantify distinct subcomponents of evidence-accumulation processes in binary decision tasks (Ratcliff & Rouder, 1998). The parameters of the drift-diffusion model and their interpretation in the context of our experiment are illustrated in Figure 17, panel A; the modeling results are shown in Figure 17, panel B. For both rule-aware (n = 14) and rule-unaware (n = 21) participants, grammar learning was manifested in non-decision parameter t_0, which captures processes that affect response times but are not actually related to information uptake from the presentation of the stimulus at 0 ms (in our case, the English noun). In other words, t_0 captures changes in response time that occur not because a decision-making process was itself brought closer to its conclusion, but rather because the timing of the entire process was shifted forward or backward due to some decision-agnostic factor. In most drift diffusion studies using trial designs with only one stimulus (e.g., a single word that is presented or a single image to categorize), the t_0 parameter is usually associated with factors of little theoretical interest like low-level perception or motor-related processes, which are not affected by the main stimulus. However, because in my trial design the nouns are preceded by pseudoword articles, in my analysis the parameter t_0 can be interpreted as capturing effects from the pre-stimulus pseudoword article which are unrelated to how information from the noun is processed. For instance, t_0 could be affected if participants anticipate the correct button press prior to the presentation of the noun due to low-level motor preparation that did not change how the noun was actually read (e.g., did not lead to accumulation of evidence from the noun at a slower/faster rate; did not change whether evidence was "pre-accumulated" before reading the noun; and did not alter the threshold of evidence necessary before providing a response). By contrast, for rule-aware participants only, we found that learning also affected parameter *z*, which is a parameter that captures bias in evidence accumulation towards or against the correct response at the beginning of each trial before the stimulus is shown (Abugaber & Morgan-Short, 2021). This significant change in parameter *z* suggests that the pseudoword article provided some information for the process of evidence accumulation that begins when the noun is presented.

In sum, my drift-diffusion modeling suggests that only the rule-aware participants in my pilot data actually derived any linguistic information from the pseudoword article that was pertinent to the livingness of the noun, whereas all groups (both rule-aware and rule-unaware) derived non-linguistic information that allowed them to initiate the response regardless of the upcoming noun. Extrapolating these to my dissertation results, I can say, at a minimum, that my paradigm allows for both linguistic learning (as suggested by effects on parameter *z* in rule-aware learners) and non-linguistic learning (as suggested by effects on parameter t_0 in both rule-aware and rule-unaware learners).

5.2.3 No learning found in trial designs without predictable buttons

Another piece of evidence suggesting that ostensive grammar learning in my experiment was driven by lower-level motor responses in my design is that, when the trial design is altered to remove stimulus-response contingencies, rule learning effects disappeared. Behavioral piloting was conducted via the Internet-based Amazon Turk Platform that counterbalanced the current trial design against two alternate versions of the experiment: one in which the response key assignment was randomized for each trial and communicated to the participant via a response cue (much like in the noun-only block of this study), and another version wherein the response key assignment was switched at the end of each epoch (i.e., after every 77 trials or so). Both modifications meant that the proportion of left vs. right key presses would be equal for each of the *ul/gi/ro/ne* artificial language pseudowords. Results showed that rule-learning effects in Blocks 1 and 2 disappeared entirely under these modifications, and in the randomized button design they did not even appear in Block 3 after the participant was explicitly told the rule (see Figure 18). This suggests that the learning effect in my experiment may be predicated on a consistent mapping between the artificial language stimuli and a fixed button response.

5.2.4 Prototypicality of noun animacy has no effect on grammar rule application

Yet another way to examine the extent to which linguistic meaning played a role in my experimental task beyond simple motor patterns is by examining whether the grammatical rule is harder to implement for nouns that are more/less typical for their living/nonliving category. To illustrate this, “bird” and “protestor” may both technically be living things, but the former specifies a biological category and thus may be more evocative of the prototypical idea of a living thing. By contrast, a word like “protestor” may be more specific to a political context and abstract in the extent to which it evokes the idea of a living thing.

To conduct such an analysis of prototypicality of animacy, I quantified the similarity of each stimulus noun to the other nouns in the same living/nonliving category using three different metrics from computational linguistics: Wu-Palmer distance (Wu & Palmer, 1994), which is based on the number of “nodes” that separate two items in the hierarchy of features in the human-annotated WordNet semantic database (Miller, 1995); *word2vec* cosine distance (Mikolov et al., 2013), which gauges the similarity of two words based on a shallow feedforward network on word co-occurrences within local contexts (i.e., focusing on neighboring words); and *GloVe* cosine distance (Pennington et al., 2014), which quantifies word similarity via matrix factorization of global word-to-word co-occurrence counts in an entire corpus. In separate regression models for each of these measures on per-trial reaction times, I found (a) a main effect of trial condition (rule-adhering vs. rule-violating); (b) a main effect of word frequency (included in the model to control for the fact that highly frequent words may co-occur more with other words by definition); and (c) a main effect of the word’s average similarity to the other words in its living/nonliving category (for Wu-Palmer and *word2vec* but not for *GloVe*). However, in no case was there a significant interaction between the noun’s rule-adhering vs. rule-violating condition and that noun’s similarity to other words in its respective living/nonliving category.

This finding that the prototypical livingness/nonlivingness of a word did not have a significant effect on how the grammatical rule was applied suggests that the rule may not have truly been dependent on identifying the living/nonliving status of a word.

5.2.5 Learning varied with running proportion of violation trials per pseudoword

Another data point that suggests that low-level mechanisms may have driven learning of the hidden grammar rule in my experiment comes from examining how performance in the experiment is affected by the running proportion of rule-adhering vs. rule-violating trials for each pseudoword in my experiment. Recall that in my design one randomly chosen trial out of every seven consecutive trials violates the hidden rule. This means that, over the course of the experiment, the proportion of rule-adhering trials for each pseudoword would vary slightly. To illustrate this, although on average roughly 13% (one-seventh) of trials for each of *gi, ro, ul,* and *ne* would be violations, this actual figure would differ at any given point in the experiment based on which pseudoword article the participants had seen more violations for. This variation becomes more pronounced when I look at the statistics within a running window of recently seen trials.

In a linear mixed model on individual trial response times with random intercepts for participant and item and the fixed factors trial condition (rule-adhering vs. rule-violating), trial number across the experiment (to account for a natural speed-up as participants become accustomed to the task), proportion of rule-adhering trials for the current pseudoword in a window of thirty preceding trials, and the interaction of trial condition and proportion of rule-adhering trials, I found a significant interaction such that, as the proportion of rule-adhering trials goes up, reaction times for rule-violating trials are slower and reaction times for rule-adhering trials are faster. By contrast, this difference between rule-violating and rule-adhering trials disappears as the proportion of rule-adhering trials in the window goes down (see Figure 19).

These results suggest that, rather than involving fixed linguistic knowledge (e.g., "*gi* goes with living things in the majority of cases") that is applied equally throughout the experiment once it is acquired, the target of learning may be an amorphous and constantly varying stimulus-response contingency such that participants constantly adapt to shifting stimulus-response pairings to different degrees based on the kinds of trials that they had seen most recently. This sort of lower-level statistical adaptation would seem rather far removed from higher-level processing of semantic concepts like “living/nonliving.”

5.2.6 Arguments against low-level learning in my experiment

Even if a low-level button-pressing contingency to some extent facilitated learning in the experiment, I argue that this is not necessarily the sole mechanism involved. To begin with, knowledge of the stimulus-response contingencies in my design would only come as the result of many trials of exposure during which participants would have to read the nouns to provide the correct living/nonliving response (as they would not have a way of surmising that the artificial language article was predictive of animacy in earlier portions of the experiment). Secondly, the reaction time effects that I have shown compare *correctly responded* rule-adhering vs. rule-violating trials. This means that, for participants to have provided a correct response on a violation trial, the participant must have, at a minimum, read the noun and taken into account its living/nonliving status. As such, the participant must have been reading and processing the animacy of the noun on some level. Thirdly, my drift-diffusion model results did find that parameter z (related to bias in an evidence-accumulation process from noun reading) was affected by trial condition in the rule-aware group, suggesting at least some interaction with word meaning rather than pure motor prediction from the preceding pseudoword. Finally, considering what Batterink et al. (2014, p. 171) call the “automatic nature of reading” and taking into account the fact that median epoch response times in my data were in the 800-1600 ms range, it is unlikely that participants were not reading and processing the noun to some degree when these were displayed. In this way, I argue that my task did indeed feature some linguistic component.

5.2.7 Final conclusion in regard to interpretation of experiment results

In sum, the points presented above suggest that learning in this experimental paradigm involves both a linguistic and a non-linguistic component. Both of these effects would lead to our observed results of strong behavioral learning effects. However, neither of these effects may have been strong enough in regard to neurocognitive processes to be detected via EEG. This would explain the pattern of results I found wherein strong behavioral effects were found but EEG findings were for the most part nonsignificant.

5.3 Relevance to prior research in psychology and SLA

5.3.1 Comparison to Batterink et al. (2014)

How do my findings compare with those of Batterink et al. (2014), the study upon which this experiment was most closely based? Table IX presents a summary of my behavioral and ERP findings in relation to Batterink et al.'s. While I replicated their behavioral findings of a main effect of Trial Condition (rule-violating vs. rule-adhering) with no interaction with Rule Awareness status (rule-aware vs. rule-unaware), unlike them I did not find the Condition x Epoch interaction in participant response times or accuracies. This may have been due to differences between the experiments as described further below. The starkest difference in results was in the ERP findings: whereas Batterink et al. found a negativity for rule-unaware participants and a positivity for rule-aware participants, the only significant effect I found was a positivity for rule-aware participants. As such, I might say that I generally replicated their behavioral results but not their ERP results.

How to explain this pattern of results? For the ERP results more specifically, as discussed above (see section 5.1.1), the failure to replicate their ERP results may be attributed to the finding that natural cross-participant ERP variability in language processing may limit the conclusions that can be drawn from comparisons of grand averaged groups (Tanner, 2019;

Tanner et al., 2018). As a broader caveat, however, another possible reason for my failure to replicate may be due in part to differences between my experiment and theirs. Namely, whereas Batterink and colleagues had a 90-minute nap between the first two blocks of the experiment, this was replaced by a 5-minute break in my experiment. This may have contributed to fatigue in my participants. Furthermore, whereas Batterink and colleagues showed both the pseudoword article and the English noun on the same screen, my trial design showed these on separate screens in sequence (to allow us to investigate prediction as per Research Question 2). This may have made the experimental task more difficult in that participants had to retain the near/far meaning of the pseudoword in memory as they read the noun instead of reading them together. Furthermore, in order to allow for a third block of trials in the experiment, and in order to balance word features like concreteness and imageability across animate/inanimate words (so as to allow for confound-free decoding of an English noun's living or nonliving status in Research Question 2), I expanded my word stimuli for the living category beyond animals to also include arguably more abstract entities like humans (with words like "professor," "student," etc.), which may have made rule learning and application more difficult. These changes between my design and that of Batterink et al. (2014) may have altered the nature of the task such that, even though behavioral effects were found, the corresponding ERPs would not appear. However, because both rule-aware and rule-unaware participant groups showed significant behavioral differences between rule-adhering and rule-violating trials, I can at least surmise that learning still occurred even in this adapted version of the artificial language task.

5.3.2 Is learning without awareness possible?

As described in my literature review, several linguistic (e.g., Williams, 2005; Leung & Williams, 2011; Batterink et al., 2014) and non-linguistic (e.g., Nissen & Bullemer, 1987; Nissen et al., 1987, 1988) studies have found evidence of learning without conscious awareness of said learning. My results support these findings in that even participants who were unable to identify the hidden rule after prompting nevertheless showed significant behavioral and (to a lesser extent) EEG signs of learning the hidden pattern in my experiment. Conversely, my results go against prior research that has failed to replicate findings of grammar learning in rule-unaware participants (Hama & Leow, 2010; Faretta-Stutenberg & Morgan-Short, 2011). By extension, these findings of grammar learning without awareness support the L2 psycholinguistic models from Tomlin and Villa (1994) and Leow (2015) and contradict the models of Schmidt (1990) and Robinson (1995) which posit that attention and awareness are both necessary for L2 acquisition to occur.

5.3.3 The interplay of implicit and explicit processing

Going beyond whether rule-unaware grammar learning is possible, what relevance do these results have for the interface between implicit and explicit grammar processing? In terms of our behavioral effects, our findings of no significant differences in behavioral performance between rule-aware and rule-unaware participants suggest that, at least in the context of incidental learning as in our experiment, explicit information does not necessarily facilitate learning (contradicting the model from DeKeyser, 2007), to the extent that the rule-aware group did not show stronger learning than the rule-unaware group. That said, because the rule-aware group did not show *weaker* learning than the rule-unaware group, we did not find that explicit processing hinders learning overall; this contrasts with findings of a blocking effect as reported in Ellis and Sagarra (2010). Under an interpretation of our results wherein all participants' learning was underlyingly driven by implicit learning, we might also surmise that explicit processes have no effect on implicit learning (supporting Paradis, 2009), because rule-aware participants did not perform any differently from rule-unaware participants as far as behavioral measures are concerned.

If we take our MVPA analysis for Research Question 2 at face value, and if we assume that rule-awareness entails explicit processing, then the fact that we did not find neural signs of implicit processing during rule-aware periods would support the “no interface” position wherein implicit and explicit grammar processing are entirely separate processes, such that explicit processing does not entail in any way that implicit processing is also occurring (Paradis, 2009). However, this absence of evidence may also be attributed to our aforementioned issues with decoding sensitivity.

If we go beyond our formal quantitative analyses of behavioral and EEG data and take a more qualitative look at the data, however, our findings yielded some indication of a negative effect of explicit processes on performance (and presumably on implicit learning), similar to reported “blocking” effects wherein explicit knowledge can hinder L2 learning outcomes (e.g., Ellis & Sagarra, 2010). First, visual inspection of my rule-unaware participants’ reaction times (Figure 8) and accuracies (Figure 11) shows that the learning effect is manifested not just in terms of faster responses for rule-adhering trials but also as a drop in performance (i.e., slower reaction times and lower accuracies) for rule-violating trials especially around epoch 9 (immediately after they had been told the hidden rule). In other words, learning the rule explicitly entails doing objectively worse on exceptions to the rule than if one did not have any explicit knowledge at all. Corroborating evidence for this is also found in the debriefing responses, which indicate a tension between task performance and conscious deliberation of the underlying pattern: one participant “tried to figure [the rule] out while answering and it would make [them] slower and forget what [they were] thinking.” Another said that they “kind of thought about it but then started focusing on speed more than anything so [they] didn't give it too much thought." Another participant "was too focused on the task to analyze this relationship [between the pseudowords and nouns]." Yet another participant "thought of [the rule] as a possibility during the experiment but during the experiment was more focused on putting in the right inputs." Tellingly, another participant said that they "tried to figure out the pattern and thought that it might be about living and nonliving but because the questions kept on coming was more focused on answering quickly." As seen in these subjective reports, explicit processing may at times be in conflict with optimal task performance (at least for rule-violating trials specifically).

Going beyond the interface debate specifically within the field of SLA, our findings may also be informative for the wider field of cognitive psychology. My behavioral findings that rule-aware and rule-unaware learners did not show significant differences in learning effects could be interpreted as evidence that grammar learning in this paradigm was underlyingly driven by implicit processes, which would support prior findings from cognitive psychology that implicit learning may operate simultaneously with explicit learning (as in Curran & Keele, 1993) and proceed equally whether or not explicit learning co-occurs (supporting Taylor, Krakauer, & Ivry, 2014). My finding that learning in my study might not necessarily depend on access to noun meaning but may instead be driven by simple non-linguistic statistical associations would concur with prior research demonstrating that learning can occur entirely through lower level mechanisms, e.g., in the context of serial reaction time task performance by participants with chance-level awareness (Willingham et al., 1989); by neuropsychological patients with limited declarative memory (Scoville & Milner, 1957; Corkin, 1968); and by participants with pharmacologically-induced transient amnesia and thus limited explicit memory (Nissen &

Bullemer, 1987; Nissen et al., 1987, 1988). Furthermore, our tentative (qualitative) findings that participants sometimes showed objectively worse performance on rule-violating trials after becoming rule-aware parallels prior findings that explicit strategies can at times hurt performance when these are applied to well-practiced sequences (Beilock & Carr, 2001; Beilock et al., 2002; Castaneda & Gray, 2007; Flegal & Anderson, 2008; Perkins-Ceccato et al., 2003; Poolton et al., 2006; Zachry et al., 2005; Tanaka & Watanabe, 2018).

Another way that my study looks at the interplay of implicit and explicit processes is by examining how these might differ in terms of linguistic prediction. Recall that our MVPA analysis yielded no evidence of semantic prediction from the artificial language pseudoword. Taken at face value, at first blush these results go against prior findings of semantic prediction in psycholinguistic studies that have used behavioral response times (e.g., Federmeier et al., 2010), eye-movement tracking (Lew-Williams & Fernald, 2007), and EEG (e.g., Lau et al., 2013). However, given the fact that learning in my study may have been (at least partly) non-linguistic (as discussed in Section 5.2), it is possible that any prediction in my study was ultimately non-linguistic, in the style of motor anticipation as in a serial reaction time task (Nissen & Bullemer, 1987). As such, it may be problematic to extrapolate my null findings to make strong claims about semantic prediction in grammar processing

Finally, my dissertation aimed to investigate the nature of implicit vs. explicit grammar processing by examining whether reaction times are time-locked with ERP indices of explicit processing. However, because we only found a significant ERP effect in rule-unaware learners our analysis was ultimately performed on ERP indices of *implicit* processing. If we assume that the late positive deflections that we observed for these learners represent P600 effects, then our conclusion that reaction times and ERP latencies were underlyingly correlated (but that this correlation was not borne out in all of my analyses due to weak ERP results) would have several implications for the interpretation of ERPs in prior literature. Firstly, my findings of a positive component in rule-unaware learners suggest that the P600 ERP component may not be limited to explicit processing. Secondly, if P600s are time-locked to response times, then P600 delays to low-frequency words (as in Allen et al., 2003) or in second language processing (Kotz, 2009) may be attributed to delays in processing of words that require more time to access (whether because they are rarer or because they are not in one's native language).

In sum, our finding of no behavioral difference between implicit vs. explicit learning disfavors models with a strong interface between implicit and explicit learning (e.g., DeKeyser, 2007), favoring instead theories with a weak or no interface (e.g., Paradis, 2009). Our MVPA results also favor theories with a weak or no interface (e.g., Paradis, 2009), though interpretation of these should be tempered by our weak observed decoding sensitivity. However, follow-up analyses of our behavioral and debriefing data suggest a negative effect of explicit processing (as in Ellis & Sagarra, 2010). We found no evidence for semantic prediction in grammar processing at the neural level (cf. Lau et al., 2013), though note again the limits with our MVPA analysis. Finally, our analyses correlating reaction times with the timing of our observed late positivity may be relevant for interpreting P600 ERP effects in prior research.

5.3.4 Relevant findings for future research using artificial language methodologies

I share here two additional observations from our data that would be relevant for future researchers using an artificial language learning methodology. Firstly, I found that the target of learning is not always clearly defined: as described above, many participants (10 out of 48) reported noticing a recurring pattern in the button press sequences in the experiment. As such, one might argue that for them the target of learning was not actually "living/nonliving" but rather a lower-level idiosyncratic feature of the experiment context (i.e., “left/right button press”). Even if they noticed the living/nonliving regularity in addition to the recurring key press patterns (as 5 of these 10 did), it would be difficult to show conclusively that these participants were basing their decisions on the livingness/nonlivingness of the nouns rather than the button press pattern. Furthermore, other participants reported noticing that certain pseudowords were often accompanied by words for animals or for professions (e.g., “teacher,” “professor,” “student,” etc.). As such, for these participants one might again argue that they did not truly learn “living/nonliving” but rather some more specific rule that merely correlated with the actual regularity (although they would have still been coded as rule-aware under our criteria). Both of these findings emphasize the fact that the purported target of a grammar rule may not actually the target of learning. Future L2 research should take this into account when designing experiments and drawing conclusions from the same.

Secondly, I found that rule-aware participants do not necessarily apply the explicit knowledge that they have. In a behavioral pilot study, I compared 40 participants who learned the hidden rule incidentally (as in the current ERP experiment) vs. 90 participants who received explicit instruction about the hidden rule from the very beginning of the experiment. Sample size differences between these conditions were accounted for by a bootstrapping procedure wherein, for each of 10,000 iterations, a subsample of 40 instructed participants was drawn to match the 40 incidental procedures and the analysis was re-run, allowing me to create a histogram of *p* values from which the stability of the results was verified. Mixed 2x2x8 ANOVAs on the between-participants factors Experiment Condition (incidental vs. instructed) and the within-participants factors Trial Type (rule-adhering vs. rule-violating) and Epoch (for each of eight epochs) did not find any differences between instructed and incidental participants in either reaction times or accuracies. However, an interesting qualitative difference emerged between instructed and incidental learners (which was probably not detected in our inferential statistics because the ANOVA-based analysis is not equipped to distinguish between a gradual vs. sudden effect of Epoch in the three-way interaction of Experiment Condition x Trial Type x Epoch): namely, incidental learners showed a large initial dip in accuracies for rule-violating trials in Block 1 but comparable accuracies between rule-adhering and rule-violating trials in Block 2. By contrast, instructed participants showed a consistent moderate drop in accuracy for rule-violating trials following a five-minute break halfway through the experiment (see Figure 20). One interpretation of this qualitative effect is that participants who learn grammar incidentally may overapply rules at first but then cease applying them when they see exceptions to the rules. By contrast, instructed participants may feel obligated to apply the grammatical rule, as they might feel that this is expected of them. This pressure may come from what Wonnacott and colleagues (2015, p. 475) call "experimental pragmatics (i.e., they realize what the experimenter is asking of them, irrespective of whether that is warranted by the input)." In this way, explicit knowledge does not necessarily entail explicit processing.

5.4 Limitations

A major issue with generalizing my results to the field of SLA is that the relative contributions of implicit vs. explicit processes to L2 attainment may depend on the exact circumstances on learning. For instance, prior research suggests that L2 instruction is most effective when it is matched to the individual learner's profile, which is affected by age, cognitive maturity, cognitive learning style, motivation, attitudes, personality, L1 background, statistical learning ability, and L2 proficiency level at time of instruction (e.g., Tagarelli et al., 2011; Misyak et al., 2012). Another factor that can affect the relative interplay of implicit vs.

explicit learning is the nature of the target L2 feature, e.g., the form's perceptual salience, complexity, frequency, redundancy, similarity with the L1, consistency/reliability of a linguistic rule or pattern, and transparency between form and function (e.g., DeKeyser, 1998; Doughty & Williams, 1998) or whether it comprises grammar, vocabulary, or pronunciation (e.g., Mackey et al., 2000; Jiang, 2004). Task demands may also affect the implicitness/explicitness of processing (e.g., DeKeyser, 2012; Robinson, 2001b), entailing different explicit strategies (see N. Ellis, 2005; Sanz & Leow, 2011) such that, just like consciousness itself may not be binary but instead involve a continuum (Baddeley, 1976, 1997; Cowan, 1998; Solso, 1998), explicit learning strategies might span a wide continuum of depth of processing (Chechile et al., 1981; Williams, 1999) which warranting different pedagogical approaches (Doughty, 2001). As De Graaff and Housen put it, "[e]ffective instruction… is context-appropriate, that is, its effect, relevance and usefulness depend on goals, learners, resources, and the environment" (p. 728, 2009).

Of course, no single study can aim to reasonably capture the effects of all of the factors described above. However, for the context of our experiment certain intervening variables might be particularly problematic. Firstly, because participants were recruited from introductory psychology classes at a university, the age is particularly skewed to 18–20-year-olds with relatively high levels of education, meaning that our generalization to younger/older learners from other educational backgrounds is quite limited. Secondly, because this was an unpaid experiment and participants were not compensated differently based on their performance, motivation might have varied widely between participants (as illustrated by one disqualified participant who provided only chance-level responses with implausibly fast reaction times throughout the experiment). Thirdly, our participants varied widely between each other in terms of the additional languages that they spoke beyond English, which potentially entails differences in the salience of an animate/inanimate distinction. In particular, 17 of the 48 participants included in the final analysis reported proficiency in Spanish, which marks animate direct objects differently than inanimate direct objects. Finally, our semi-artificial language comprises just four words with a pattern that was design 87% consistent (i.e., for six out of every seven instances), meaning that our results might not speak to grammatical forms with more surface-level variety and lower/higher consistency.

Another limitation to generalizing my findings is that my dissertation only compares implicit, rule-unaware processing with explicit, rule-aware processing that comes after a period of rule-unaware processing—and *not* to explicit, rule-aware processing that comes as the result of direct instruction at an initial phase. This might not reflect learning in classroom settings that provide explicit grammar instruction at the early stages of (if not entirely before) exposure to L2 forms, especially in light of order effects wherein L2 forms that are acquired at an earlier stage can “block” the learning of subsequent forms due to competition for attention (Ellis, 2006; Ellis & Sagarra, 2010; Solman & Chung, 1996).

Additional limitations include the fact that our experiment is a single-session laboratory study and thus might not give enough time/exposure to yield robust learning effects in EEG. Furthermore, using participant self-reports to assess awareness of the hidden grammatical rule is admittedly imperfect because participants might underreport or overreport explicit knowledge (see Leow & Hama, 2013). Finally, combining a known language (English) with an artificial language comprising only four novel pseudowords represents an extremely simplified version of learning a natural language. To my knowledge, the ecological validity of such semi-artificial language designs has never been examined empirically.

Given these limitations, the results of this dissertation study would represent only an initial step towards exploring the interplay of implicit vs. explicit processing in L2 morphosyntax acquisition. Because this might play out differently based on an assortment of factors when taken to the "real world," any findings would need to be validated with actual learners before being put into practice.

5.5 Conclusion

The replication and extension of Batterink et al. (2014) presented here aims to advance our understanding of the interplay between implicit and explicit processing in second language morphosyntax learning. In terms of L2 psycholinguistic theory, my reproduction of previous findings of implicit L2 morphosyntax learning—in terms of both behavioral reaction time effects (as in Leung & Williams, 2012) and the finding of event-related potentials in rule-unaware learners (as in Batterink et al., 2014, though note the difference in effect polarity)— supports models of second language acquisition that predict that implicit learning is possible (Tomlin & Villa, 1994; Leow, 2015) over models that do not (Schmidt, 1990, 1995, 2001; Robinson, 1995, 2003). My finding of no evidence from MVPA that implicit processing co-occurs during periods of rule-aware processing speaks to the “interface debate” in the field of second language acquisition by supporting frameworks that posit a weak interface (e.g., N. Ellis, 2005) or no interface (Krashen, 1981; Paradis, 1994) between implicit and explicit processing, over frameworks that posit a strong interface (e.g., DeKeyser, 1997), though this may be due to limited MVPA decoding sensitivity. My MVPA analyses on semantic prediction yielded no neural evidence of semantic prediction in either rule-aware or rule-unaware learners (cf. Lau et al., 2013), though again this may be due to low MVPA sensitivity. Finally, our analyses correlating reaction times with the timing of our observed late positivity show that this ERP effect is closely linked to external responses and may be relevant for interpreting P600 ERP effects in prior research.

Findings from this research on the relative role of implicit vs. explicit learning in SLA could have tangible takeaways for L2 praxis, in as much as language teachers, curriculum designers, and students themselves have the ability to manipulate attention and conscious processing towards specific L2 features. To illustrate this, my finding that explicit processing of L2 forms has no drawbacks for concurrent implicit learning might suggest that one would have little to lose from adopting a teaching approach that includes metalinguistic instruction and other explicit methods (at least as far as implicit acquisition is concerned). That said, my participants' subjective reports and observed accuracies also suggest that the allocation of attention to particular features of the L2 input may work against implicit learning in certain cases, such that it may be more beneficial to prioritize one style of teaching over the other at times. In this way, beyond simply informing second language psycholinguistic theory, the findings from this dissertation could have meaningful implications for L2 praxis.

CITED LITERATURE

Abugaber, D. & Morgan-Short, K. (2021). Differences between implicit vs. explicit grammar learning as revealed by drift diffusion modeling. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society 2021*. https://cutt.ly/BmQyAL

Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: event-related potentials reflect continuous model updates. *eNeuro, 4*(6).

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes, 18*(4), 405-430.

Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing: Evidence from differential object marking acquisition. *Studies in Second Language Acquisition, 37*(2), 237-268.

Andringa, S., & Rebuschat, (2015). New directions in the study of implicit and explicit learning: An introduction. *Studies in Second Language Acquisition, 37*(2), 185-196.

Anguera, J. A., Reuter-Lorenz, A., Willingham, D. T., & Seidler, R. D. (2010). Contributions of spatial working memory to visuomotor learning. *Journal of Cognitive Neuroscience, 22*(9), 1917-1930.

Baddeley, A. (1976). *The psychology of memory*. New York: Basic Books.

Baddeley, A. (1997). *Human memory: Theory and practice (Rev. ed.)*. Hove: Psychology Press.

Barušs, I. (1987). Metanalysis of definitions of consciousness. *Imagination, Cognition and Personality, 6*(4), 321-329.

Batterink, L. J., Karns, C. M., & Neville, H. (2011). Dissociable mechanisms supporting awareness: the P300 and gamma in a linguistic attentional blink task. *Cerebral Cortex, 22*(12), 2733-2744.

Batterink, L. J., & Neville, H. J. (2013). The human brain processes syntax in the absence of conscious awareness. *Journal of Neuroscience, 33*(19), 8528-8533.

Batterink, L. J., Oudiette, D., Reber, J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia, 65*, 169-179.

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory, 22*(11), 544-556.

Battista, J. R. (1978). The science of consciousness. In K. Pope, & J. Singer (Eds.), *The stream of consciousness* (pp. 55-87). Springer, Boston, MA.

Beilock, S. & Carr, T. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*, 701–725.

Beilock, S., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, *8*(1), 6.

Bell, & Collins, L. (2009). \`It's vocabulary\`/\`It's gender\`: learner awareness and incidental learning. *Language Awareness, 18*, 277-293

Benati, A. (2004). The effects of processing instruction and its components on the acquisition of gender agreement in Italian. *Language Awareness*, *13*(2), 67-80.

Bennett, C. M., Miller, M. B., & Wolford, G. L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *Neuroimage, 47*(Suppl 1), S125.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues.* Hove, UK: Erlbaum

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*, 69-83.

Bialystok, E. (1990). The competence of processing: Classifying theories of second language acquisition. *TESOL Quarterly*, *24*(4), 635-648.

Bond, K., Gabriele, A., Fiorentino, R., & Aleman Banon, J. (2011). Individual differences and the role of the L1 in L2 processing: An ERP investigation. In J. Herschensohn & D. Tanner (Eds.), *Proceedings of the 11th generative approaches to second language acquisition conference (GASLA 2011)* (pp. 17-29). Somerville, MA: Cascadilla Proceedings Project.

Bornkessel-Schlesewsky I. & Schlesewsky, M. (2009). *Processing syntax and morphology: A neurocognitive perspective, Vol. 6*. Oxford University Press.

Bowers, K. S., & Meichenbaum, D. (1984). *The unconscious reconsidered*. New York: Wiley.

Bowles, M. (2010). *The think-aloud controversy in second language research*. New York: Routledge.

Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1-42). Oxford, England: Lawrence Erlbaum.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology*, *64*(1), 74-95.

Brouwer, S., & Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *Journal of Psycholinguistic Research, 45*(5), 1151-1160.

Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400 and P600 in language comprehension. *Frontiers in Psychology, 8*, 1327.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.

Caldwell-Harris, C. L., Lancaster, A., Ladd, D. R., Dediu, D., & Christiansen, M. H. (2015). Factors influencing sensitivity to lexical tone in an artificial language: Implications for second language learning. *Studies in Second Language Acquisition, 37*(2), 335-357.

Carlson, T. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*, 1-17.

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1-1.

Castaneda, B., & Gray, R. (2007). Effects of focus of attention on baseball batting performance in players of differing skill levels. *Journal of Sport and Exercise Psychology*, *29*(1), 60-77.

Cauchoix, M., Arslan, A. B., Fize, D. & Serre, T. (2012). The neural dynamics of visual processing in monkey extrastriate cortex: A comparison between univariate and multivariate techniques. In M. Grosse-Wentrup, B. Murphy, & I. Rish (Eds.), *Machine learning and interpretation in neuroimaging* (pp. 164-171). Springer, Berlin, Heidelberg.

Cauchoix, M., Barragan-Jason, G., Serre, T. & Barbeau, E. J. (2014). The neural dynamics of face detection in the wild revealed by MVPA. *Journal of Neuroscience, 34*(3), 846-854.

Chechile, R., Richman, C., Topinka, C., & Ehrensbeck, K. (1981). A developmental study of the storage and retrieval of information. *Child Development, 52,* 251-278.

Choi, J. T., Bouyer, L. J., & Nielsen, J. B. (2014). Disruption of locomotor adaptation with repetitive transcranial magnetic stimulation over the motor cortex. *Cerebral Cortex*, *25*(7), 1981-1986.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(1), 1-15.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* Greenwood Publishing Group.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,* 170-178.

Cintrón-Valentín, M., & Ellis, N. C. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition*, *37*(2), 197-235.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2*, 406–416.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Combs, C. H. (2004). What cognitive processes are triggered by input enhancement? *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, 4*, 1-14.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia, 6*(3), 255-265.

Correia, J. M., Jansma, B., Hausfeld, L., Kikkert, S., & Bonte, M. (2015). EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations. F*rontiers in Psychology*, *6*, 1-10.

Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition*, *21*(2), 162-167.

Cowan, N. (1998). *Attention and memory: An integrated framework* (Vol. 26). Oxford University Press.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* Irvington.

Cunningham, H. A. (1989). Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 493-506.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 189.

Dawson, M. E., & Schell, A. M. (1987). Human autonomic and skeletal classical conditioning: The role of conscious cognitive factors. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 27-55). Oxford, England: John Wiley & Sons.

De Graaff, R. (1997). The Experanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*(2), 249-276.

De Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 726-755). Oxford, UK: Blackwell.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*(3), 379-410.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*(2), 195-221.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In C. J. Doughty & H. M. Long (Eds.), *The handbook of second language acquisition* (pp. 312–348). Oxford, UK: Blackwell.

DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Lawrence Erlbaum.

DeKeyser, R. (2009). Cognitive-psychological processes in second language learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 119-138). Oxford, UK: Blackwell.

DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62*, 189-200.

DeKeyser, R., & Juffs, A. (2005). Cognitive considerations in L2 learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 461-478). Routledge.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21.

Delorme, A., Westerfield, M., & Makeig, S. (2007). Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. *Journal of Neuroscience*, *27*(44), 11949-11959.

Destrebecqz, A, & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin and Review*, *8,* 343-350.

Dienes, Z. (2012). Conscious versus unconscious learning of structure. In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 337-364). Berlin, Germany: De Gruyter Mouton.

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, *4*(1), 3-23.

Dienes, Z., & Longuet-Higgins, C. (2004). Can musical transformations be implicitly learned? *Cognitive Science*, *28*, 531– 558.

Doughty, C. J. (2001). The cognitive underpinnings of focus on form. In Robinson (Ed.), *Cognition and second language instruction* (pp. 206-257). Cambridge: Cambridge University Press.

Doughty, C. J. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. H. Long (Eds.), The handbook of second language acquisition (p256-310). Malden, MA: Blackwell.

Doughty, C. J., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., & Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *European Journal of Neuroscience*, *8*(4), 637-648.

Dussias, E., Valdes-Kroff, J. R., Guzzardo-Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition, 35*, 353–387.

Eichenbaum, H. (2011). *The cognitive neuroscience of memory: An introduction.* Oxford University Press.

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*(2), 305-352.

Ellis, N. (2006) Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164-194.

Ellis, N. (2007). SLA: The associative-cognitive creed. In B. VanPatten and J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77-96). Cambridge: Cambridge University Press.

Ellis, N., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, *32*(4), 553-580.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, *27*, 91–113.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79-114). Ann Arbor, MI: Academic Press.

Ellis, R. (1997). *SLA research and language teaching.* New York: Oxford University Press.

Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition*, *24*(2), 223-236.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*(2), 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*(2), 141-172.

Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339-360). Oxford; New York: Oxford University Press.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing, and teaching*. Bristol, UK: Multilingual Matters.

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

Ewolds, H. E., Bröker, L., De Oliveira, R. F., Raab, M., & Künzell, S. (2017). Implicit and explicit knowledge both improve dual task performance in a continuous pursuit tracking task. *Frontiers in Psychology, 8*, 2241.

Fahrenfort, J. J., Grubert, A., Olivers, C. N., & Eimer, M. (2017). Multivariate EEG analyses support high-resolution tracking of feature-based attentional selection. *Scientific Reports*, *7*(1), 1886.

Fahrenfort, J. J., Van Driel, J., Van Gaal, S. & Olivers, C. N. (2018). From ERPs to MVPA using the Amsterdam decoding and modeling toolbox (ADAM). *Frontiers in Neuroscience*, *12*, 368.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, *115*(3), 149-161.

Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition, 30*(3), 277-305.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499-504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 458.

Fitts, M., & Posner, M. I. (1967). *Human Performance*. Brooks. Cole, Belmont, CA.

Flegal, K. E., & Anderson, M. C. (2008). Overthinking skilled motor performance: Or why those who teach can't do. *Psychonomic Bulletin & Review*, *15*(5), 927-932.

Foster, J. J., Bsales, E. M., Jaffe, R. J., & Awh, E. (2017). Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Current Biology, 27*(20), 3216-3223.

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology, 115*(1), 168-177.

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science, 7*(6), 585-594.

Freud, S. (1915). Das Unbewusste. *Internationale Zeitschrift für Ärztliche Psychoanalyse*, *3*(4), 189-203.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*(3), 183-192.

Galea, J.M., Vazquez, A., Pasricha, N., de Xivry, J.J. & Celnik, (2011) Dissociating the roles of the cerebellum and motor cortex during adaptive learning: The motor cortex retains what the cerebellum learns. *Cerebral Cortex, 21,* 1761–1770.

Gallistel, C. R. (1990). *The organization of learning.* The MIT Press.

Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition, 38*(2), 177-215.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35*(3), 483-517.

Godfroid, A., & Schmidtke, J. (2013). What do eye movements tell us about awareness? A triangulation of eye-movement data, verbal reports and vocabulary learning scores. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 183–205). Honolulu: University of Hawaii, National Foreign Language Resource Center.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443-482). John Benjamins.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neuroscience*, *18*(22), 9420-9428.

Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, *13*, 168–184.

Gregg, K. R. (1984). Krashen's monitor and Occam's razor. *Applied Linguistics, 5*(2), 79-100.

Gregory, E., McCloskey, M., Ovans, Z., & Landau, B. (2016). Declarative memory and skill-related knowledge: Evidence from a case study of amnesia and implications for theories of memory. *Cognitive Neuropsychology, 33*(3-4), 220-240.

Grey, S., Tanner, D., & van Hell, J. G. (2017). How right is left? Handedness modulates neural responses during morphosyntactic processing. *Brain Research, 1669*, 27-43.

Grey, S., Williams, J. N., & Rebuschat, (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition, 36*(4), 611-645.

Grey, S., Williams, J. N., & Rebuschat, (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences, 38*, 44-53.

Grootswagers, T., Cichy, R., & Carlson, T. (2016). Predicting behavior from decoded searchlight representations shows where decodable information relates to behavior. *Perception, 45*, 360.

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience, 29*(4), 677-697.

Groppe, D. M., Urbach, T. and Kutas, M. (2011), Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*, 1711–1725.

Görgen, K., Hebart, M. N., Allefeld, C., & Haynes, J. D. (2018). The same analysis approach: Practical protection against the pitfalls of novel neuroimaging analysis methods. *Neuroimage*, *180*, 19-30.

Haith, A. M. & Krakauer, J. W. (2013). Model-based and model-free mechanisms of human motor learning. In M. Richardson, M. Riley, & K. Shockley (Eds.), *Progress in motor control* (pp. 1-21). Springer, New York.

Hama, M., & Leow, R. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition, 32*(3), 465-491.

Hamrick, P. (2014). Recognition memory for novel syntactic structures. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 68*(1), 2-7.

Hamrick, P., & Rebuschat, (2012). How implicit is statistical learning? In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365-382). Berlin, Germany: De Gruyter Mouton.

Hamrick, P., & Rebuschat, (2014). Frequency effects, learning conditions, and the development of implicit and explicit lexical knowledge. In J. Connor-Linton & L. Amoroso (Eds.), *Measured language: Quantitative approaches to acquisition, assessment, processing, and variation* (pp. 125-139). Georgetown University Press: Washington, D.C

Hamrick, P., & Sachs, R. (2018). Establishing evidence of learning in experiments employing artificial linguistic systems. *Studies in Second Language Acquisition, 40*(1), 153-169.

Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics, 10*(3), 331– 358.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425-2430.

Haynes, J. D., & Rees, G. (2006). Neuroimaging: Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*(7), 523.

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain: A Journal of Neurology, 120*(1), 123-140.

Hebart, M. (2016). *Multivariate pattern analysis and brain decoding.* Functional Magnetic Resonance Imaging Core Facility summer course lecture, July 8, 2018. Accessed online at https://fmrif.nimh.nih.gov/public/fmri-course/fmri-course-summer-2016

Heikel, E., Sassenhagen, J., & Fiebach, C. J. (2018a). Time-generalized multivariate analysis of EEG responses reveals a cascading architecture of semantic mismatch processing. *Brain and Language, 184*, 43-53.

Heikel, E., Sassenhagen, J., & Fiebach, C. J. (2018b). Decoding semantic predictions from EEG prior to word onset. *bioRxiv*, 393066.

Heuer, H., & Hegele, M. (2008). Adaptation to visuomotor rotations in younger and older adults. *Psychology and Aging, 23*(1), 190.

Housen, A., Pierrard, M., & Van Daele, S. (2005). Structure complexity and the efficacy of explicit grammar instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in Instructed Second Language Acquisition* (pp. 235-270). Berlin/New York: Mouton de Gruyter.

Hulstijn, J. & Ellis, R. (Eds.). (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning [Special issue]. *Studies in Second Language Acquisition, 27*(2).

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research, 18*(3), 193-223.

Hulstijn, J. (2015) Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning: The virtues and pitfalls of a two-system view. In Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 25-46). John Benjamin Publishing Company.

Hulstijn, J., & De Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *Aila Review, 11*(97-112).

Hulstijn, J., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning, 34*(1), 23-43.

Hwang, E. J., Smith, M. A., & Shadmehr, R. (2006). Dissociable effects of the implicit and explicit memory systems on learning control of reaching. *Experimental Brain Research, 173*(3), 425-437.

Imamizu, H., & Shimojo, S. (1995). The locus of visual-motor learning at the task or manipulator level: Implications from intermanual transfer. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 719.

Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology, 7*(3), e1002012.

Jamalabadi, H. (2017). *Optimizing parameters and algorithms of multivariate pattern classification for hypothesis testing in high-density EEG.* Doctoral dissertation, Eberhard Karls Universität Tübingen.

Jamalabadi, H., Alizadeh, S., Schonauer, M., Leibold, C., & Gais, S. (2016). Classification based hypothesis testing in neuroscience: Below-chance level classification rates and overlooked statistical properties of linear parametric classifiers. *Human Brain Mapping, 37*, 1842-55.

Jamalabadi, H., Alizadeh, S., Schönauer, M., Leibold, C., & Gais, S. (2018). Multivariate classification of neuroimaging data with nested subclasses: Biased accuracy and implications for hypothesis testing. *PLoS Computational Biology, 14*(9), e1006486.

James, W. (1890). *The principles of psychology, Vol. 2*. New York: Henry Holt and Company.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. In G. Deuschl & A. Eisen (Eds.), *Recommendations for the Practice of Clinical Neurophysiology: Guidelines of the International Federation of Clinical Physiology (EEG Suppl. 52)* (pp. 370-375). Elsevier Science B.V.

Jian, Y., & Chun, M. M. (2003). Contextual cueing: Reciprocal influences between attention and implicit learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 277-296). Amsterdam: Benjamins.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603-634.

Jiménez, L., Vaquero, J. M. K., & Lupianez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 475-490.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping, 14*(3), 166-185.

Kahneman, D., & Egan, (2011). *Thinking, fast and slow (Vol. 1)*. New York: Farrar, Straus and Giroux.

Kalatzis, I., Piliouras, N., Ventouras, E., Papageorgiou, C. C., Rabavilas, A. D., & Cavouras, D. (2004). Design and implementation of an SVM-based computer classification system for discriminating depressive patients from healthy controls using the P600 component of ERP signals. *Computer Methods and Programs in Biomedicine, 75*(1), 11-22.

Kanfer, R., & Ackerman, L. (1989). Dynamics of skill acquisition: Building a bridge between intelligence and motivation. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 83–136). Mahwah, NJ: Erlbaum.

Kaplan, J. T., & Meyer, K. (2012). Multivariate pattern analysis reveals common neural patterns across individuals during touch observation. *Neuroimage*, *60*(1), 204-212.

Kaplan, J. T., Man, K., & Greening, S. G. (2015). Multivariate cross-classification: applying machine learning techniques to characterize abstraction in neural representations. *Frontiers in Human Neuroscience, 9*, 151.

Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage*, *225*, 117465.

Keele, S.W., Ivry, R.B., Mayr, U., Hazeltine, E. & Heuer, H. (2003) The cognitive and neural architecture of sequence representation. *Psychology Reviews, 110*, 316–339

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology, 51*(1), 1-21.

Kelly, M.H. & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua, 92*, 105-140.

Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology, 45*(2), 250-274.

Kihlstrom, J. F. (1984). Conscious, subconscious, unconscious: A cognitive perspective. In K. S. Bowers & D. Meichenbaum (Eds.), *The unconscious reconsidered* (pp. 149-211). New York: Wiley.

Kim, A. E., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: An ERP investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(3), 406.

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences, 18*(4), 203-210.

King, J. R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two distinct dynamic modes subtend the detection of unexpected sounds. *PloS one*, *9*(1), e85791.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399-1402.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*(2-3), 68-74.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal, 13*, 8-17.

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience, 20*(23), 8916-8924.

Krashen, S. (1977). The Monitor Model for adult second language performance. In M. Burt, H. Dulay, & M. Finocchiaro (Eds.), *Viewpoints on English as a second language: In honor of James E. Alatis* (pp. 152-161). New York, NY: Regents.

Krashen, S. (1979). The Monitor Model for second language acquisition. In R. Gingras (Ed.), *Second language acquisition and foreign language teaching* (pp. 1-26). Arlington, VA: Center for Applied Linguistics.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford, UK: Pergamon.

Krashen, S. (1983). Newmark's "Ignorance Hypothesis" and current second language acquisition theory. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 135-153). Rowley, MA: Newbury House.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley Longman Ltd.

Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London, UK: Academic Press.

Krashen, S., & Terrell, T.D. (1983). *The natural approach: Language acquisition in the classroom*. London: Prentice Hall Europe.

Kriegeskorte N., Simmons W. K., Bellgowan S. F., & Baker C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12*, 535–540.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23-49

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621-647.

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

Lamal, A. (1990). On the importance of replication. *Journal of Social Behavior and Personality, 5*(4), 31.

Lambert, A. J., & Sumich, A. L. (1996). Spatial orienting controlled without awareness: A semantically-based implicit learning effect. *Quarterly Journal of Experimental Psychology, 49*, 490-518.

Lambert, T. (2002). Visual orienting, learning and conscious awareness. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 253-275). Amsterdam: Benjamins.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Laszlo, S., & Federmeier, K. D. (2014). Never seem to find the time: evaluating the physiological time course of visual word recognition with regression analysis of single-item event-related potentials. *Language, Cognition and Neuroscience*, *29*(5), 642-661.

Lau, E. F., Holcomb, J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience, 25*(3), 484-502.

Lavecchia, A. (2015). Machine-learning approaches in drug discovery: Methods and applications. *Drug Discovery Today, 20*(3), 318-331.

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87-118.

Leow, R. & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition, 26*, 35-58.

Leow, R. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis. *Applied Psycholinguistics, 19*, 133–159.

Leow, R. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22*, 557-584

Leow, R. (2015). *Explicit learning in the L2 classroom*. New York: Routledge.

Leow, R. (2018). ISLA: How implicit or how explicit should it be? Theoretical, empirical, and pedagogical/curricular issues. *Language Teaching Research*, *1*, 1-18.

Leow, R., & Martin, A. (2017). Enhancing the input to promote salience of the L2: A critical overview. In *Salience in second language acquisition* (pp. 167-186). Routledge.

Leow, R., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning*, *13*(2), 1-16.

Leung, J. H., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually derived meanings. *Studies in Second Language Acquisition*, *33*(1), 33-55.

Leung, J. H., & Williams, J. N. (2012). Constraints on implicit learning of grammatical form-meaning connections. *Language Learning, 62*(2), 634-662.

Leung, J. H., & Williams, J. N. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition, 36*(4), 733-755.

Lewicki, P. (1986). *Nonconscious social information processing*. New York: Academic Press.

Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology, 17*(1), 26-65.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193-198.

Li, Z., Meier, M. A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters, 45*(10), 4773-4779.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.

Lopez-Calderon, J., & Luck, S. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, 213.

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

Luck, S. (2018, September 16). On the application of decoding/classification/MVPA approaches to ERP data. *ERP Info blog*. https://erpinfo.org/blog/2018/9/16/decoding

Luck, S., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146-157.

Luck, S., Vogel, E., & Shapiro, K. (1996) Word meanings can be accessed but not reported during the attentional blink. *Nature, 383*(6601), 616-618.

Lunzer, E. A. (1979). The development of consciousness. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (pp. 1-19). New York: Academic Press.

Lyons, W. E. (1986). *The disappearance of introspection.* Boston: MIT Press.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*(4), 471-497.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163.

Malone, L. A., & Bastian, A. J. (2010). Thinking about walking: Effects of conscious correction versus distraction on locomotor adaptation. *Journal of Neurophysiology, 103*(4), 1954-1962.

Mandler, G. (1975). Consciousness: Respectable, useful, and probably necessary. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J.: Erlbaum.

Marasco, D. E., & Kontokosta, C. E. (2016). Applications of machine learning methods to identifying and predicting building retrofit opportunities. *Energy and Buildings, 128*, 431-441.

Marathe, A., Ries, A.J., & McDowell, K. (2013). A novel method for single-trial classification in the face of temporal variability. In D.D. Schmorrow, C.M. Fidopiastis (Eds.), *Augmented Cognition* (pp. 345-352). Heidelberg: Springer.

Marcel, J. A. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197–237

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning, 68*(2), 321-391.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1083-1100.

Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of Neuroscience, 26*(14), 3642-3645.

McLaughlin, B. (1990). Conscious versus unconscious learning. *TESOL Quarterly, 24*, 617–634.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language learning, 33*(2), 135-158.

Meara, M. (2005). *LLAMA language aptitude tests*. Swansea, UK: Lognostics.

Michel, C., Pisella, L., Prablanc, C., Rode, G., & Rossetti, Y. (2007). Enhancing visuomotor adaptation by reducing error signals: single-step (aware) versus multiple-step (unaware) exposure to wedge prisms. *Journal of Cognitive Neuroscience, 19*(2), 341-350.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In Y. Bengio & Y. LeCun (Eds.), *Proceedings of Workshop at International Conference on Learning Representations 2013*. Scottsdale, Arizona, USA, May 2-4, 2013,. Available via http://arxiv.org/abs/1301.3781

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38(*11), 39-41.

Miller, J. (1988). A warning about median reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 539

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage, 53*(1), 103-118.

Misyak, J. B., Goldstein, M. H., & Christiansen, M. H. (2012). Statistical-sequential learning in development. In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 13-54). Berlin, Germany: De Gruyter Mouton.

Mizumori, S. J. (2016). Self-regulation of memory processing centers of the brain. In P. Jackson, A. Chiba, R. Berman, & M. Ragozzino, *The Neurobiological Basis of Memory* (pp. 199-225). Springer, Cham.

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement: A reply to Tanner et al. C*ortex, 66*, 156-159.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences, 14*(10), 435-440.

Morgan-Short, K. (2014). Electrophysiological approaches to understanding second language acquisition: A field reaching its potential. *Annual Review of Applied Linguistics, 34,* 15-36.

Morgan-Short, K. (2020). Insights into the neural mechanisms of becoming bilingual: A brief synthesis of second language research with artificial linguistic systems. *Bilingualism: Language and Cognition, 23*(1), 87-91.

Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P. C., & Wong, F. C. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, *37*(2), 383-419.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, C. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition, 17*(1), 56-72.

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.

Morgan-Short, K., & Wood-Bowden, H. W. (2006). Processing instruction and meaningful output-based instruction: Effects on second language development. *Studies in Second Language Acquisition, 28*(1), 31-65.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432-434.

Naccache, L., Marti, S., Sitt, J. D., Trübutschek, D., & Berkovitch, L. (2016). Why the P3b is still a plausible correlate of conscious access? A commentary on Silverstein et al., 2015. *Cortex*. Elsevier, 2016, 10.1016/j.cortex.2016.04.003. hal-01317650

Natsoulis, T. (1987). The six basic concepts of consciousness and William James' stream of thought. *Imagination, Cognition and Personality, 6*, 289-319.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*(2), 151-165.

Nikolić, D., Häusler, S., Singer, W., & Maass, W. (2009). Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS biology*, *7*(12), e1000260.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than I can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231.

Nissen, M. J., & Bullemer, (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1-32.

Nissen M. J., Knopman, D. S., & Schacter, D.L. (1987). Neurochemical dissociation of memory systems. *Neurology, 37*(5), 789-94.

Nissen, M. J., Ross, J. L., Willingham, D. B., Mackenzie, T. B., & Schacter, D. L. (1988). Memory and awareness in a patient with multiple personality disorder. *Brain and Cognition, 8*(1), 117-134.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 21(5), 1209.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528.

Oakley, D. (1985). Animal awareness, consciousness, and self-image. In D. Oakley (Ed.), *Brain and Mind*. New York: Methuen.

Odlin, T. (1986). On the nature and use of explicit knowledge. *IRAL-International Review of Applied Linguistics in Language Teaching, 24*(1-4), 123-144.

O'Keefe, J. (1985). Is consciousness the gateway to the hippocampal cognitive map? A speculative essay on the neural basis of mind. In D. Oakley (Ed.), *Brain and Mind*. New York: Methuen.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods, 8*(4), 434.

O'Malley, J. M., Chamot, A. U., & Walker, C. (1987). Some applications of cognitive theory to second language acquisition. *Studies in Second Language Acquisition, 9*(3), 287-306.

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

Paciorek, A., & Williams, J. N. (2015). Implicit learning of semantic preferences of verbs. *Studies in Second Language Acquisition, 37*(2), 359-382.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam, John Benjamins.

Park, E. S. (2005). Constraints of implicit focus on form: Insights from a study of input enhancement. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, 4*, 1–30.

Payne, B. R., Lee, C. L., & Federmeier, K. D. (2015). Revisiting the incremental effects of context on word processing: Evidence from single-word event-related brain potentials. *Psychophysiology*, *52*(11), 1456-1469.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1532–1543. Available via https://www.aclweb.org/anthology/D14-1162

Perkins-Ceccato, N., Passmore, S. R., & Lee, T. D. (2003). Effects of focus of attention depend on golfers' skill. *Journal of Sports Sciences, 21*(8), 593-600.

Pernet, C. R., Wilcox, R. R., & Rousselet, G. A. (2013). Robust correlation analyses: false positive and power validation using a new open source MATLAB toolbox. *Frontiers in Psychology*, *606*(3).

Perruchet, P. (2008). Implicit learning. In J. Byrne & H. L. Roediger (Eds.), *Learning and memory: A comprehensive reference* (Vol. 2) (pp. 597– 621). Oxford, UK: Elsevier.

Perruchet, P., & Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(4), 785.

Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, *25*(1), 99-126.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). The ICLabel dataset of electroencephalographic (EEG) independent component (IC) features. *Data in Brief*, *25*, 104101.

Plessner, H., Betsch, C., & Betsch, T. (Eds.). (2008). *Intuition in judgment and decision making.* New York: Erlbaum.

Poli, R., Cinel, C., Citi, L., & Sepulveda, F. (2010). Reaction-time binning: A simple method for increasing the resolving power of ERP averages. *Psychophysiology*, *47*(3), 467-485

Poolton, J. M., Maxwell, J. P., Masters, R. S. W., & Raab, M. (2006). Benefits of an external focus of attention: Common coding or conscious processing? *Journal of Sports Sciences, 24*(1), 89-99.

Posner, M. I. (1992). Attention as a cognitive and neural system. *Current Directions in Psychological Science, 1*, 11–14.

Posner, M. I. (1994). Attention: The mechanism of consciousness. *Proceedings of the National Academy of Sciences, 91*, 7398–7403.

Posner, M. I. (1995). Interaction of arousal and selection in the posterior attention network. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 390–405). London: Clarendon.

Posner, M. I., & Dehaene, S. (2000). Attentional networks. In M. Gazzaniga (Ed.), *Cognitive Neuroscience: A reader* (pp. 156-170). Oxford: Blackwell.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25-42.

Posner, M. I., Petersen, S. E., Fox, T., & Raicle, M. E. (1988). Localization of cognitive functions in the human brain. *Science, 240*, 1627–1631.

Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A.D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91–112). Cambridge, MA: MIT Press.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psych. Science, 9*(5), 347–356.

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., ... & Kosslyn, S. M. (1995). A PET investigation of implicit and explicit sequence learning. *Human Brain Mapping, 3*(4), 271-286.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 317– 327.

Reber, A. S. (1989). Implicit learning of tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219– 235.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious.* Oxford, UK: Oxford University Press

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

Rebuschat, (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning, 63*(3), 595-626.

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition, 37*(2), 299-334.

Rebuschat, P., Hamrick, P., Sachs, R., Riestenberg, K., & Ziegler, N. (2013). Implicit and explicit knowledge of form-meaning connections: Evidence from subjective measures of awareness. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 255–275). Honolulu: University of Hawaii, National Foreign Language Resource Center.

Rebuschat, P., & Williams, J. (2009). Implicit learning of word order. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 31*, 31.

Rebuschat, P., & Williams, J. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics, 33*(4), 829-856.

Redding, G. M., & Wallace, B. (1988). Components of prism adaptation in terminal and concurrent exposure: Organization of the eye-hand coordination loop. *Perception & Psychophysics*, *44*(1), 59-68.

Redding, G. M., & Wallace, B. (1993). Adaptive coordination and alignment of eye and hand. *Journal of Motor Behavior, 25*(2), 75-88.

Redding, G. M., & Wallace, B. (1996). Adaptive spatial alignment and strategic perceptual-motor control. *Journal of Experimental Psychology: Human Perception and Performance, 22*(2), 379.

Redding, G. M., & Wallace, B. (1997). Prism adaptation during target pointing from visible and nonvisible starting locations. *Journal of Motor Behavior, 29*(2), 119-130.

Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language learning, 45*(2), 283-331.

Robinson, P. (1996a). *Consciousness, rules, and instructed second language acquisition*. New York, NY: PeterLang.

Robinson, P. (1996b). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18,* 27–67.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty, & M. H. Long (Eds.) *The handbook of second language acquisition* (pp. 631-678). Malden, MA: Blackwell Publishing.

Robinson, P. (2001a). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27-57.

Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press

Robinson, P. (2010). Implicit artificial grammar and incidental natural second language learning: How comparable are they? *Language Learning, 60*, 245-263.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043-1056.

Rogers, J., Revesz, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional morphology. *Applied Psycholinguistics, 37*(4), 781-812.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness & Cognition, 24*, 214– 222.

Romberg, A. R. & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 906-914.

Rosa, E. M., & Leow, R. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics, 25*(2), 269-292.

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters, 384*(3), 228-233.

Roth, W. T., Ford, J. M., & Kopell, B. S. (1978). Long-latency evoked potentials and reaction time. *Psychophysiology, 15*(1), 17-23.

Rothbart, M., & Posner, M. (2001). Mechanism and variation in the development of attentional networks. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 353-363). Cambridge, MA: MIT Press.

Rousselet, G. A., & Wilcox, R. R. (2020) Reaction times and other skewed distributions: problems with the mean and the median. *Meta-Psychology, 4*, MP.2019.16.

Rutherford, E. W. & Sharwood-Smith, S. M. (1985). Consciousness-raising and universal grammar, *Applied Linguistics*, *6*(3), 274–282.

Rutiku, R., & Bachmann, T. (2017). Juxtaposing the real-time unfolding of subjective experience and ERP neuromarker dynamics. *Consciousness and Cognition, 54*, 3-19.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*(1), 172-196.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27-52.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*(2), 101-105.

Sanz, C., & Leow, R. (Eds.) (2011). *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism.* Washington: Georgetown University Press.

Sanz, C., & Morgan-Short, K. (2005). Explicitness in pedagogical interventions: Input, practice, and feedback. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 234-263). Washington: Georgetown University 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.

Schelonka, K., Graulty, C., Canseco-Gonzalez, E., & Pitts, M. A. (2017). ERP signatures of conscious and unconscious word and letter perception in an inattentional blindness paradigm. *Consciousness and Cognition, 54*, 56-71.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schimidt (Ed.), *Attention and awareness in foreign language learning (technical report #9)* (pp. 1-63). Honolulu, Hawaii: University of Hawaii, Second Language Teaching & Curriculum Center.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC 2010, Singapore, December 2-4* (pp 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language learning* (pp. 237-322). Rowley, MA: Newbury House.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition, 15*(2), 147-163.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*(1), 11.

Seidler, R. D., Purushotham, A., Kim, S. G., Ugurbil, K., Willingham, D., & Ashe, J. (2005). Neural correlates of encoding and expression in implicit sequence learning. *Experimental Brain Research*, *165*(1), 114-124.

Seliger, H. W. (1979). On the nature and function of language rules in language teaching. *TESOL Quarterly*, *13*(3), 359-369.

Seliger, H. W. (1983). The language learner as linguist: Of metaphors and realities. *Applied linguistics*, *4*(3), 179-191.

Shanks, D. R. (2005). Implicit learning. In K. Lambert & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202– 220). London: Sage.

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., & Krakauer, J. W. (2012). Overcoming motor "forgetting" through reinforcement of learned actions. *Journal of Neuroscience*, *32*(42), 14617-14621a.

Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, *23*(1), 103-124.

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics*, *18*, 268-286.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS biology*, *4*(6), e179.

Solman, R. T., & Chung, K. K. (1996). Language transfer and blocking in second language vocabulary learning. *Journal of Behavioral Education*, *6*(2), 173-190.

Solso, R. L. (1998). *Cognitive psychology* (5th ed.). Boston, MA: Allyn & Bacon

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, *60*(2), 263-308.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171-177.

Squire, L. R. (2009). Memory and brain systems: 1969–2009. *Journal of Neuroscience*, *29*(41), 12711-12716.

Stefanucci, L., Buffalo, E. A., Schmolck, H., & Squire, L. R. (2000). Profound amnesia after damage to the medial temporal lobe: A neuroanatomical and neuropsychological profile of patient EP. *Journal of Neuroscience*, *20*(18), 7024-7036.

Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language aptitude testing. In Robinson (Ed.), *Individual differences and instructed language learning* (pp. 13-44). New York: John Benjamins.

Su, L., Fonteneau, E., Marslen-Wilson, W., & Kriegeskorte, N. (2012, July). Spatiotemporal searchlight representational similarity analysis in EMEG source space. In *Proceedings of the 2nd International Workshop on Pattern Recognition in NeuroImaging* (pp. 97-100). London: IEEE.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics, 38*(5), 1229-1261.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning, 65*, 860–895.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning, 67*(4), 747-790.

Szewczyk, J. M., & Schriefers, H. (2011). Is animacy special?: ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research, 1368*, 208-221.

Tagarelli, K., Ruiz, S., Vega, J. L. M., & Rebuschat, (2016). Variability in second language learning: The roles of individual differences, learning conditions, and linguistic complexity. *Studies in Second Language Acquisition, 38*(2), 293-316.

Tagarelli, K., Borges Mota, M., & Rebuschat, (2011). The role of working memory in implicit and explicit language learning. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 2061–2066). Austin, TX: Cognitive Science Society.

Tanaka, K., & Watanabe, K. (2017). Explicit instruction of rules interferes with visuomotor skill transfer. *Experimental Brain Research, 235*(6), 1689-1700.

Tanner, D., Goldshtein, M., & Weissman, B. (2018). Individual differences in the real-time neural dynamics of language comprehension. *Psychology of Learning and Motivation, 68*, 299-335.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition, 17*(2), 277-293.

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology, 52*(8), 997-1009.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289-301.

Taylor, J. A., & Ivry, R. (2011). Flexible strategies during motor learning. *PLoS Computational Biology, 7(3)*, e10001096

Taylor, J. A., & Ivry, R. (2012). The role of strategies in motor learning. *Annals of the New York Academy of Sciences, 1251*(1), 1-12.

Taylor, J. A., & Ivry, R. (2013). Implicit and explicit processes in motor learning. In W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 63-87). Cambridge: MIT Press.

Taylor, J. A., & Thoroughman, K. A. (2007). Divided attention impairs human motor adaptation but not feedback control. *Journal of Neurophysiology, 98*(1), 317-326.

Taylor, J. A., & Thoroughman, K. A. (2008). Motor adaptation scaled by the difficulty of a secondary cognitive task. *PLoS One, 3(*6), e2485.

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *Journal of Neuroscience, 34*(8), 3023-3032.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-203.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327– 369.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne, 26*(1), 1-12.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*(4), 552.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934-1945.

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology, 38*(5), 816-827.

Vafaee, P., Suzuki, Y., & Kachinske, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*, 59–95. doi:10.1017/S0272263115000455

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour, 124*, 203-220.

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*, 287–301

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*(2), 225-243.

VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

VanPatten, B., & Williams, J. (2007). Early theories in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 17-35). Cambridge: Cambridge University Press.

Wan, L., Dienes, Z., Su, I. F., & Fu, X. (2010). *EEG correlates of conscious versus unconscious knowledge in artificial grammar learning.* 14th Meeting of the Association for the Scientific Study of Consciousness (ASSC), Toronto, Canada, 24-27 June 2010.

Warden, M., Lapkin, S., Swain, M. & Hart, D. (1995). Adolescent language learners on a three-month exchange: Evidence from their diaries. *Foreign Language Annals, 28,* 537-549.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. Journal of Memory and Language, *50*(1), 1-25.

Werner, S., & Bock, O. (2007). Effects of variable practice and declarative knowledge on sensorimotor adaptation to rotated visual feedback. *Experimental Brain Research, 178*(4), 554-559.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York, NY: John Wiley.

Wickens, C. D. (1989). Attention and skilled performance. In H. D. Holding (Ed.), *Human skills* (pp. 63–102). New York, NY: Academic Press.

Wickens, T. D. (2002). *Elementary Signal Detection Theory.* Oxford: Oxford University Press.

Williams, J. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition, 21*, 1-48.

Williams, J. (2004). Implicit learning of form-meaning connections. In B. VanPatten, J. Williams, S. Rott & M. Overstreet (Eds.), *Form Meaning Connections in Second Language Acquisition* (pp. 203-218). New York: Routledge.

Williams, J. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.

Williams, J. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319– 353). Bingley, UK: Emerald Press.

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155). Cambridge: Cambridge University Press.

Willingham, D.B., Nissen, M.J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition 15*(6),1047-60

Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*(2), 17-46.

Wong, W. (2004). Processing instruction in French: The roles of explicit information and structured input. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 187-205). Mahwah, NJ: Lawrence Erlbaum.

Wonnacott, E., Boyd, J. K., Thomson, J., & Goldberg, A. E. (2012). Input effects on the acquisition of a novel phrasal construction in 5-year-olds. *Journal of Memory and Language, 66*(3), 458-478.

Wu, Z. & Palmer, M. (1994). Verb semantics and lexical selection. In *Proceedings of the 32nd annual meeting of the Association for Computational Linguistics* (pp. 1-42). Las Cruces, New Mexico.

Yang, Y. & Lisberger, S.G. (2014) Role of plasticity at different sites across the time course of cerebellar motor learning. *Journal of Neuroscience, 34*, 7077–7090.

Zachry, T., Wulf, G., Mercer, J., & Bezodis, N. (2005). Increased movement accuracy and reduced EMG activity as the result of adopting an external focus of attention. *Brain Research Bulletin, 67*(4), 304-309.

Zhang, Y., Zhang, J., & Min, B. (2012). Neural dynamics of animacy processing in language comprehension: ERP evidence from the interpretation of classifier–noun combinations. *Brain and language, 120*(3), 321-331.

Table I

ANIMACY AND DISTANCE ASSIGNMENT OF THE FOUR ARTIFICIAL LANGUAGE ARTICLES FOR THE EXPERIMENT.

	Participants are not told…	
	Animate	Inanimate
Participants are told…		
Near	gi	ro
Far	ul	ne

Table II

SUMMARY STATISTICS OF PARTICIPANT DEMOGRAPHIC INFORMATION FOR BEHAVIORAL PILOTING STUDY.

Attribute	*M* (*SD*)
Gender	32 female, 20 male
Age	18.71 (1.29)
Self-reported English reading proficiency (max 10)	9.54 (0.69)
Self-reported English listening proficiency (max 10)	9.60 (0.71)
Self-reported English speaking proficiency (max 10)	9.56 (0.72)
Percent reporting additional language	69.23%
Additional language reading proficiency (max 10)	4.51 (3.39)
Additional language listening proficiency (max 10)	6.65 (3.00)
Additional language speaking proficiency (max 10)	5.74 (2.80)
Additional reported languages	Arabic (4), Cantonese (1), Chinese (1), German (1), Gujarati (4), Hebrew (1), Hmong (1), Malayalam (1), Russian (1), Spanish (17), Tagalog (1), Telugu (1), Twi (1), Urdu (2), Vietnamese (2)

Table III

ANOVA RESULTS FOR MEDIAN REACTION TIMES IN BLOCKS 1 AND 2.

	df	MSE	*F*	η^2_G	*p*
Awareness	1, 46	0.85	0.03	.001	.864
Condition	1, 46	0.03	15.17	.008	.001*
Awareness * Condition	1, 46	0.03	2.07	.001	.157
Epoch	3.67, 168.89	0.10	54.99	.238	<.001*
Awareness * Epoch	3.67, 168.89	0.10	0.31	.002	.854
Condition * Epoch	5.35, 246.12	0.03	0.97	.002	.441
Awareness * Condition * Epoch	5.35, 246.12	0.03	0.31	.001	.915

* $p < 0.05$

Table IV

ANOVA RESULTS FOR MEDIAN REACTION TIMES ON BLOCK 3 ONLY.

	df	MSE	*F*	η^2_G	*p*
Awareness	1, 36	0.94	0.07	.001	.799
Condition	1, 36	0.04	11.70	.008	.002*
Awareness * Condition	1, 36	0.04	1.79	.001	.189
Epoch	3.54, 127.31	0.10	40.95	.215	<.001*
Awareness * Epoch	3.54, 127.31	0.10	0.54	.004	.684
Condition * Epoch	5.28, 190.19	0.03	0.99	.003	.428
Awareness * Condition * Epoch	5.28, 190.19	0.03	0.49	.001	.797

* $p < 0.05$

Table V

ANOVA TABLE FOR ANALYSIS OF MEAN EPOCH ACCURACIES IN BLOCKS 1 AND 2.

	df	MSE	*F*	η^2_G	*p*
Awareness	1, 46	0.08	0.42	.004	.518
Condition	1, 46	0.01	5.34	.005	.025*
Awareness * Condition	1, 46	0.01	0.16	.001	.694
Epoch	5.51, 253.31	0.01	1.17	.007	.326
Awareness * Epoch	5.51, 253.31	0.01	0.70	.004	.637
Condition * Epoch	5.87, 269.85	0.01	1.79	.009	.104
Awareness * Condition * Epoch	5.87, 269.85	0.01	0.44	.002	.847

* $p < 0.05$

Table VI

ANOVA TABLE FOR ANALYSIS OF MEAN EPOCH ACCURACIES IN BLOCK 3.

	df	MSE	*F*	η^2_G	*p*
Awareness	1, 36	0.09	0.00	.001	.959
Condition	1, 36	0.05	31.41	.154	.001*
Awareness * Condition	1, 36	0.05	0.29	.002	.596
Epoch	2.65, 95.26	0.01	0.87	.004	.448
Awareness * Epoch	2.65, 95.26	0.01	0.30	.001	.799
Condition * Epoch	2.54, 91.32	0.02	0.83	.004	.465
Awareness * Condition * Epoch	2.54, 91.32	0.02	0.53	.003	.632

* $p < 0.05$

Table VII

RESULTS OF MIDLINE ANOVA FOR BLOCKS 1, 2, AND 3.

Block 1

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Condition	1, 46	12.78	0.58	.013	.448	1, 46	10.91	0.8	.017	.375
Awareness	1, 46	175.73	0.08	.002	.782	1, 46	278.39	1.25	.027	.269
Electrode	1.42, 65.17	12.43	5.71 *	.110	.011	1.39, 63.72	16.25	13.24 ***	.223	<.001
Awareness :Condition	1, 46	12.78	0.68	.015	.415	1, 46	10.91	0.77	.016	.386
Electrode: Condition	1.63, 75.11	0.62	0.17	.004	.804	1.57, 72.40	0.99	0.45	.010	.595
Awareness :Electrode	1.42, 65.17	12.43	0.25	.005	.705	1.39, 63.72	16.25	0.04	<.001	.904
Awareness :Electrode: Condition	1.63, 75.11	0.62	0.49	.011	.576	1.57, 72.40	0.99	0.45	.010	.594

Block 2

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Condition	1, 46	9.03	1.49	.031	.229	1, 46	14.91	1.90	.040	.175
Awareness	1, 46	155.52	0.01	<.001	.923	1, 46	222.35	0.36	.008	.551
Electrode	1.39, 63.85	12.88	2.17	.045	.138	1.50, 68.83	12.13	6.15 **	.118	.007
Awareness :Condition	1, 46	9.03	0.96	.020	.333	1, 46	14.91	4.07 *	.081	.050

Electrode: Condition	1.50, 68.87	0.90	0.46	.010	.577	1.43, 65.80	0.88	2.53	.052	.104
Awareness :Electrode	1.39, 63.85	12.88	1.65	.035	.205	1.50, 68.83	12.13	1.44	.030	.244
Awareness :Electrode: Condition	1.50, 68.87	0.90	1.39	.029	.253	1.43, 65.80	0.88	0.82	.017	.409

Block 3

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Condition	1, 46	11.72	0.60	.013	.442	1, 46	24.85	0.42	.009	.521
Awareness	1, 46	135.49	0.01	<.001	.916	1, 46	161.43	0.82	.018	.369
Electrode	1.40, 64.44	8.82	3.41 +	.069	.055	1.49, 68.70	8.12	11.18 ***	.195	<.001
Awareness :Condition	1, 46	11.72	0.14	.003	.714	1, 46	24.85	1.44	.030	.237
Electrode: Condition	1.32, 60.87	1.20	0.57	.012	.499	1.57, 72.29	1.24	0.24	.005	.731
Awareness :Electrode	1.40, 64.44	8.82	2.27	.047	.127	1.49, 68.70	8.12	1.11	.024	.321
Awareness :Electrode: Condition	1.32, 60.87	1.20	0.06	.001	.872	1.57, 72.29	1.24	0.50	.011	.566

Table VIII

RESULTS OF LATERAL ANOVA ON BLOCKS 1, 2, AND 3.

Block 1

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Awareness	1, 46	631.38	0.00	<.001	.996	1, 46	1016.02	0.47	.010	.495
Condition	1, 46	60.61	0.28	.006	.601	1, 46	54.44	0.74	.016	.393
Awareness: Condition	1, 46	60.61	0.35	.008	.558	1, 46	54.44	0.61	.013	.437
Hemisphere	1, 46	40.95	10.75 **	.189	.002	1, 46	48.77	0.00	<.001	.967
Awareness: Hemisphere	1, 46	40.95	0.66	.014	.419	1, 46	48.77	1.26	.027	.267
Laterality	1, 46	53.93	12.65 ***	.216	<.001	1, 46	74.68	14.59 ***	.241	<.001
Awareness: Laterality	1, 46	53.93	0.27	.006	.604	1, 46	74.68	2.20	.046	.144
Anteriority	1.73, 79.70	33.10	3.64 *	.073	.037	1.85, 85.28	41.88	6.49 **	.124	.003
Awareness: Anteriority	1.73, 79.70	33.10	0.29	.006	.715	1.85, 85.28	41.88	0.65	.014	.514
Condition:Hemisphere	1, 46	2.91	2.96 +	.061	.092	1, 46	6.88	2.25	.047	.140
Awareness: Condition:Hemisphere	1, 46	2.91	4.06 *	.081	.050	1, 46	6.88	2.20	.046	.145
Condition:Laterality	1, 46	4.44	0.62	.013	.434	1, 46	4.01	0.20	.004	.660
Awareness: Condition:Laterality	1, 46	4.44	1.18	.025	.284	1, 46	4.01	0.20	.004	.655

Hemisphere:Laterality	1, 46	5.54	14.08 ***	.234	<.001	1, 46	8.20	0.40	.009	.528
Awareness:Hemisphere:Laterality	1, 46	5.54	0.32	.007	.572	1, 46	8.20	0.95	.020	.335
Condition:Anteriority	1.93, 88.90	3.48	0.19	.004	.823	2.15, 99.07	2.73	0.31	.007	.751
Awareness:Condition:Anteriority	1.93, 88.90	3.48	1.25	.027	.290	2.15, 99.07	2.73	0.27	.006	.778
Hemisphere:Anteriority	2.65, 121.98	3.68	18.27 ***	.284	<.001	2.22, 102.29	7.64	20.02 ***	.303	<.001
Awareness:Hemisphere:Anteriority	2.65, 121.98	3.68	1.17	.025	.322	2.22, 102.29	7.64	0.79	.017	.470
Laterality:Anteriority	2.61, 119.91	4.62	11.17 ***	.195	<.001	2.52, 115.69	8.16	11.29 ***	.197	<.001
Awareness:Laterality:Anteriority	2.61, 119.91	4.62	1.03	.022	.376	2.52, 115.69	8.16	1.21	.026	.308
Condition:Hemisphere:Laterality	1, 46	2.50	0.03	<.001	.865	1, 46	2.74	0.12	.003	.728
Awareness:Condition:Hemisphere:Laterality	1, 46	2.50	1.42	.030	.239	1, 46	2.74	2.27	.047	.138
Condition:Hemisphere:Anteriority	1.58, 72.47	2.31	0.51	.011	.559	1.83, 84.19	2.54	1.07	.023	.344
Awareness:Condition:Hemisphere:Anteriority	1.58, 72.47	2.31	2.98 +	.061	.069	1.83, 84.19	2.54	3.08 +	.063	.056

Condition:Laterality:Anteriority	1.79, 82.33	1.09	0.66	.014	.505	1.89, 86.80	1.19	0.41	.009	.655
Awareness: Condition:Laterality:Anteriority	1.79, 82.33	1.09	2.28	.047	.114	1.89, 86.80	1.19	2.23	.046	.117
Hemisphere: Laterality:Anteriority	2.88, 132.48	2.41	2.50 +	.052	.064	2.34, 107.68	4.84	3.72 *	.075	.021
Awareness: Hemisphere: Laterality:Anteriority	2.88, 132.48	2.41	1.98	.041	.123	2.34, 107.68	4.84	1.28	.027	.283
Condition:Hemisphere:Laterality:Anteriority	1.18, 54.23	4.50	0.68	.015	.437	1.31, 60.08	4.05	0.54	.012	.512
Awareness: Condition:Hemisphere:Laterality:Anteriority	1.18, 54.23	4.50	1.42	.030	.244	1.31, 60.08	4.05	1.64	.034	.208

Block 2

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Awareness	1, 46	563.27	0.30	.007	.585	1, 46	820.11	0.00	<.001	.994
Condition	1, 46	39.62	1.49	.031	.229	1, 46	79.81	2.01	.042	.163
Awareness: Condition	1, 46	39.62	1.67	.035	.203	1, 46	79.81	4.01 +	.080	.051
Hemisphere	1, 46	46.26	7.17 *	.135	.010	1, 46	52.05	0.05	.001	.819
Awareness: Hemisphere	1, 46	46.26	0.01	<.001	.923	1, 46	52.05	1.28	.027	.264

Laterality	1, 46	51.64	9.05 **	.164	.004	1, 46	69.35	4.10 *	.082	.049
Awareness: Laterality	1, 46	51.64	0.77	.016	.386	1, 46	69.35	2.58	.053	.115
Anteriority	1.66, 76.59	41.68	1.47	.031	.237	1.84, 84.81	38.42	2.12	.044	.130
Awareness: Anteriority	1.66, 76.59	41.68	0.37	.008	.654	1.84, 84.81	38.42	0.30	.006	.726
Condition:Hemisphere	1, 46	4.60	1.31	.028	.259	1, 46	7.10	4.39 *	.087	.042
Awareness: Condition:Hemisphere	1, 46	4.60	0.07	.002	.787	1, 46	7.10	0.16	.004	.689
Condition:Laterality	1, 46	2.35	0.00	<.001	.989	1, 46	2.95	0.01	<.001	.929
Awareness: Condition:Laterality	1, 46	2.35	0.10	.002	.749	1, 46	2.95	1.54	.032	.222
Hemisphere: Laterality	1, 46	4.48	14.48 ***	.239	<.001	1, 46	8.29	0.64	.014	.427
Awareness: Hemisphere: Laterality	1, 46	4.48	0.47	.010	.495	1, 46	8.29	0.00	<.001	.964
Condition:Anteriority	1.70, 78.38	2.48	0.74	.016	.458	2.28, 104.72	1.98	0.81	.017	.461
Awareness: Condition:Anteriority	1.70, 78.38	2.48	0.47	.010	.597	2.28, 104.72	1.98	1.51	.032	.223
Hemisphere: Anteriority	2.19, 100.80	4.96	13.71 ***	.230	<.001	2.32, 106.84	6.98	14.53 ***	.240	<.001
Awareness: Hemisphere: Anteriority	2.19, 100.80	4.96	2.82 +	.058	.060	2.32, 106.84	6.98	1.05	.022	.360

Laterality:Anteriority	2.78, 128.03	4.94	6.43 ***	.123	<.001	2.41, 110.75	7.94	5.69 **	.110	.003
Awareness: Laterality:Anteriority	2.78, 128.03	4.94	1.40	.030	.248	2.41, 110.75	7.94	1.46	.031	.234
Condition:Hemisphere:Laterality	1, 46	0.67	2.90 +	.059	.095	1, 46	0.90	0.90	.019	.347
Awareness: Condition:Hemisphere:Laterality	1, 46	0.67	0.40	.009	.533	1, 46	0.90	0.01	<.001	.937
Condition:Hemisphere:Anteriority	2.41, 110.92	0.91	0.73	.016	.508	2.53, 116.31	1.18	1.59	.033	.202
Awareness: Condition:Hemisphere:Anteriority	2.41, 110.92	0.91	0.57	.012	.599	2.53, 116.31	1.18	0.38	.008	.736
Condition:Laterality:Anteriority	2.70, 124.05	0.88	2.08	.043	.113	2.90, 133.23	1.04	1.94	.041	.128
Awareness: Condition:Laterality:Anteriority	2.70, 124.05	0.88	0.70	.015	.541	2.90, 133.23	1.04	0.24	.005	.864
Hemisphere: Laterality:Anteriority	3.20, 147.11	1.91	3.01 *	.061	.029	3.44, 158.25	2.84	2.62 *	.054	.045
Awareness: Hemisphere: Laterality:Anteriority	3.20, 147.11	1.91	2.16 +	.045	.091	3.44, 158.25	2.84	1.86	.039	.130
Condition:Hemisphere:Laterality:Anteriority	3.10, 142.78	0.54	0.09	.002	.967	3.22, 148.13	0.69	0.26	.006	.870

Block 3

	400 to 800 ms					800 to 1100 ms				
	df	MSE	*F*	η^2_G	*p*	df	MSE	*F*	η^2_G	*p*
Awareness	1, 46	519.06	0.49	.011	.487	1, 46	651.03	0.02	<.001	.898
Condition	1, 46	46.59	0.35	.008	.555	1, 46	112.37	0.40	.009	.528
Awareness: Condition	1, 46	46.59	0.21	.005	.648	1, 46	112.37	0.97	.021	.331
Hemisphere	1, 46	34.43	5.21 *	.102	.027	1, 46	34.26	2.34	.048	.133
Awareness: Hemisphere	1, 46	34.43	0.08	.002	.774	1, 46	34.26	0.03	<.001	.862
Laterality	1, 46	53.24	5.11 *	.100	.029	1, 46	62.28	3.54 +	.071	.066
Awareness: Laterality	1, 46	53.24	1.59	.033	.214	1, 46	62.28	4.12 *	.082	.048
Anteriority	1.92, 88.45	30.67	1.22	.026	.299	2.12, 97.33	31.28	2.00	.042	.138
Awareness: Anteriority	1.92, 88.45	30.67	1.03	.022	.358	2.12, 97.33	31.28	0.76	.016	.477
Condition:Hemisphere	1, 46	4.40	0.00	<.001	.992	1, 46	5.99	1.60	.034	.212
Awareness: Condition:Hemisphere	1, 46	4.40	3.42 +	.069	.071	1, 46	5.99	4.22 *	.084	.046
Condition:Laterality	1, 46	4.09	1.21	.026	.278	1, 46	8.71	0.04	<.001	.842
Awareness: Condition:Laterality	1, 46	4.09	0.04	<.001	.834	1, 46	8.71	1.14	.024	.292
Hemisphere: Laterality	1, 46	4.55	10.27 **	.182	.002	1, 46	8.15	0.62	.013	.433

Awareness: Hemisphere: Laterality	1, 46	4.55	0.24	.005	.625	1, 46	8.15	0.36	.008	.550
Condition:Anteriority	1.80, 82.82	2.81	0.19	.004	.801	1.62, 74.58	5.73	0.20	.004	.772
Awareness: Condition:Anteriority	1.80, 82.82	2.81	1.16	.025	.315	1.62, 74.58	5.73	0.32	.007	.683
Hemisphere: Anteriority	2.09, 96.17	5.16	14.46 ***	.239	<.001	1.93, 88.72	8.32	18.51 ***	.287	<.001
Awareness: Hemisphere: Anteriority	2.09, 96.17	5.16	2.24	.046	.109	1.93, 88.72	8.32	1.07	.023	.345
Laterality:Anteriority	2.35, 108.26	6.33	8.76 ***	.160	<.001	2.05, 94.50	11.91	6.06 **	.116	.003
Awareness: Laterality:Anteriority	2.35, 108.26	6.33	1.25	.026	.294	2.05, 94.50	11.91	1.38	.029	.258
Condition:Hemisphere:Laterality	1, 46	0.86	0.06	.001	.816	1, 46	1.18	0.02	<.001	.884
Awareness: Condition:Hemisphere:Laterality	1, 46	0.86	0.76	.016	.389	1, 46	1.18	1.02	.022	.318
Condition:Hemisphere:Anteriority	3.22, 148.15	0.51	0.19	.004	.916	2.95, 135.48	1.23	1.92	.040	.130
Awareness: Condition:Hemisphere:Anteriority	3.22, 148.15	0.51	0.74	.016	.537	2.95, 135.48	1.23	0.54	.012	.652
Condition:Laterality:Anteriority	3.43, 157.94	0.57	0.87	.019	.469	3.23, 148.73	0.92	1.54	.032	.204

Awareness:Condition:Laterality:Anteriority	3.43, 157.94	0.57	0.95	.020	.426	3.23, 148.73	0.92	1.70	.036	.166
Hemisphere:Laterality:Anteriority	2.22, 102.15	4.15	3.33 *	.068	.035	1.83, 83.98	8.58	3.61 *	.073	.035
Awareness:Hemisphere:Laterality:Anteriority	2.22, 102.15	4.15	1.40	.029	.252	1.83, 83.98	8.58	0.73	.016	.472
Condition:Hemisphere:Laterality:Anteriority	3.38, 155.33	0.42	0.83	.018	.492	3.16, 145.16	0.66	1.49	.031	.217

Table IX

COMPARISON OF RESULTS FROM BATTERINK ET AL. (2014) WITH RESULTS OF PILOT STUDY.

Effect	Measure	Batterink et al. (2014)	Current study
Main effect of condition?	Response times	Yes	Yes
	Accuracy	Yes	Yes
	ERP	Negativity in 400-800 ms time window	No
Interaction of Condition x awareness?	Response times	No	No
	Accuracy	Stronger effect for aware vs. unaware; significant in each group separately	No
	ERP	In 400-800 ms time window, negativity only significant for unaware; in 800-1100 ms time window, positivity for aware and negativity for unaware	In 800-1100 ms time window, negativity for unaware group
Interaction of Condition x Epoch?	Response times	Yes	No
	Accuracy	Yes	No
	ERP	[Epoch not included as factor in analysis]	
Interaction of Condition x Awareness x Epoch?	Response times	No	No
	Accuracy	No	No
	ERP	[Epoch not included as factor in analysis]	

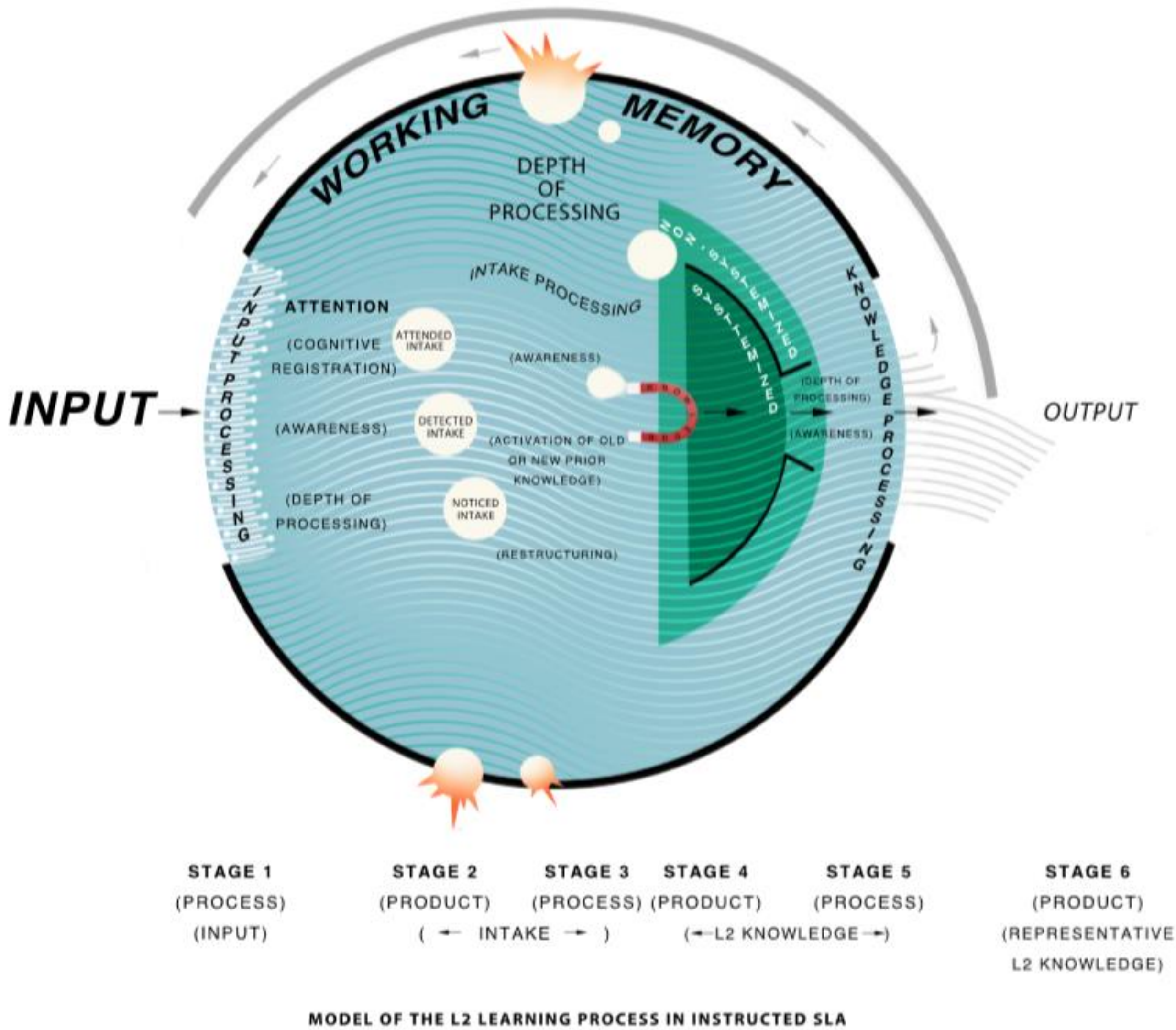


Figure 1. Model of attention in second language acquisition from Leow (2015). Reproduced with permission (see Appendix F).

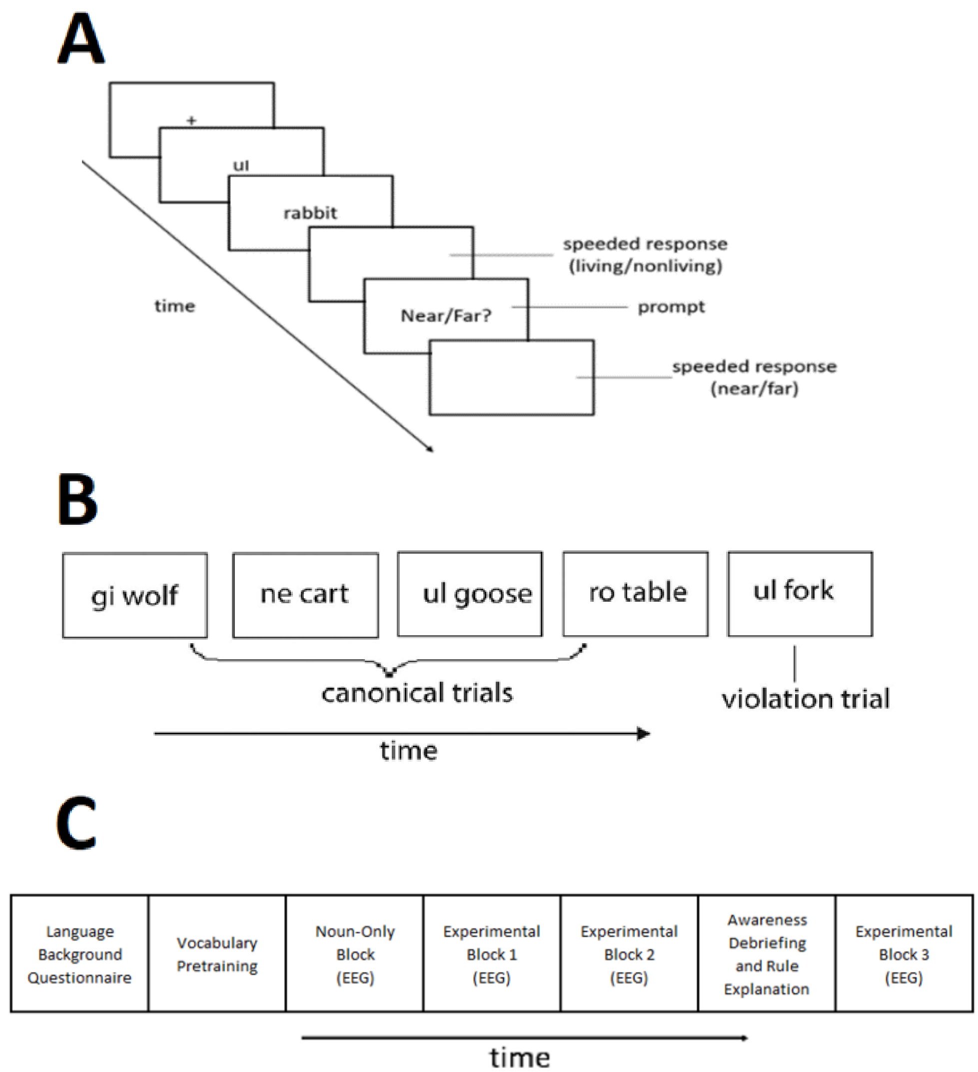


Figure 2. Summary of experimental task and overall paradigm for the proposed dissertation study, with a) sequence of events in each experimental trial, b) trial structure in the experimental task (with one out of every seven trials violating the underlying pattern, interspersed randomly), and c) overall experiment procedure. Each block comprises 308 experimental trials, with a short break between Blocks 1 and 2 and a rule awareness debriefing between Blocks 2 and 3.

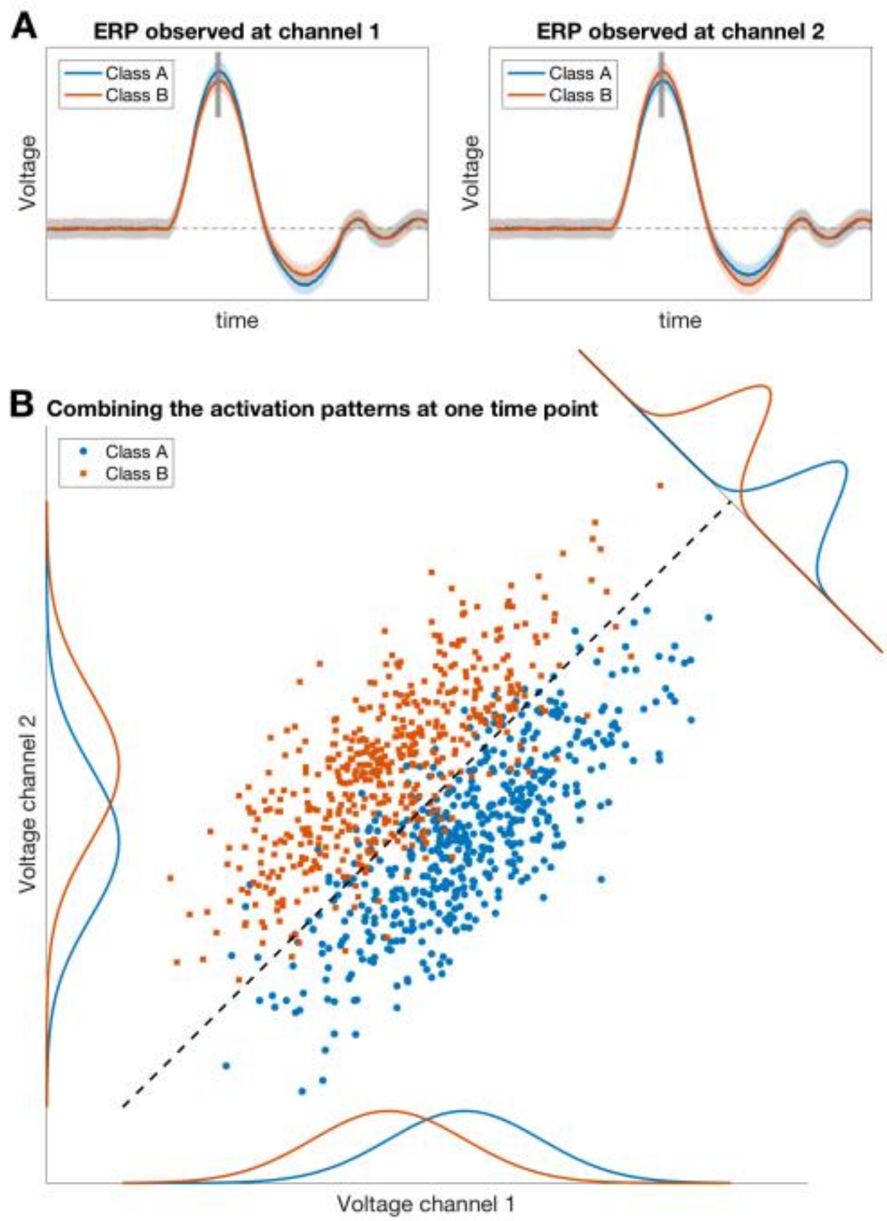


Figure 3. Illustration of the increased sensitivity of Multivariate Pattern Analysis over univariate analyses (reproduced with permission from Grootswagers et al., 2017; see Appendix F).

Panel A shows averaged event-related potentials (ERPs) in response to stimuli from two conditions (class A and class B), shown for two different channels (left and right plots). Substantial overlap between the two classes in each of the channels means that differences are not significant in a univariate analysis.

Panel B shows these same data points plotted in two-dimensional space, for a single time point in the trial (corresponding to the gray bar in the ERP plots in Panel A). The dashed line indicates the boundary that best separates the two classes, as per the underlying correlation between data points.

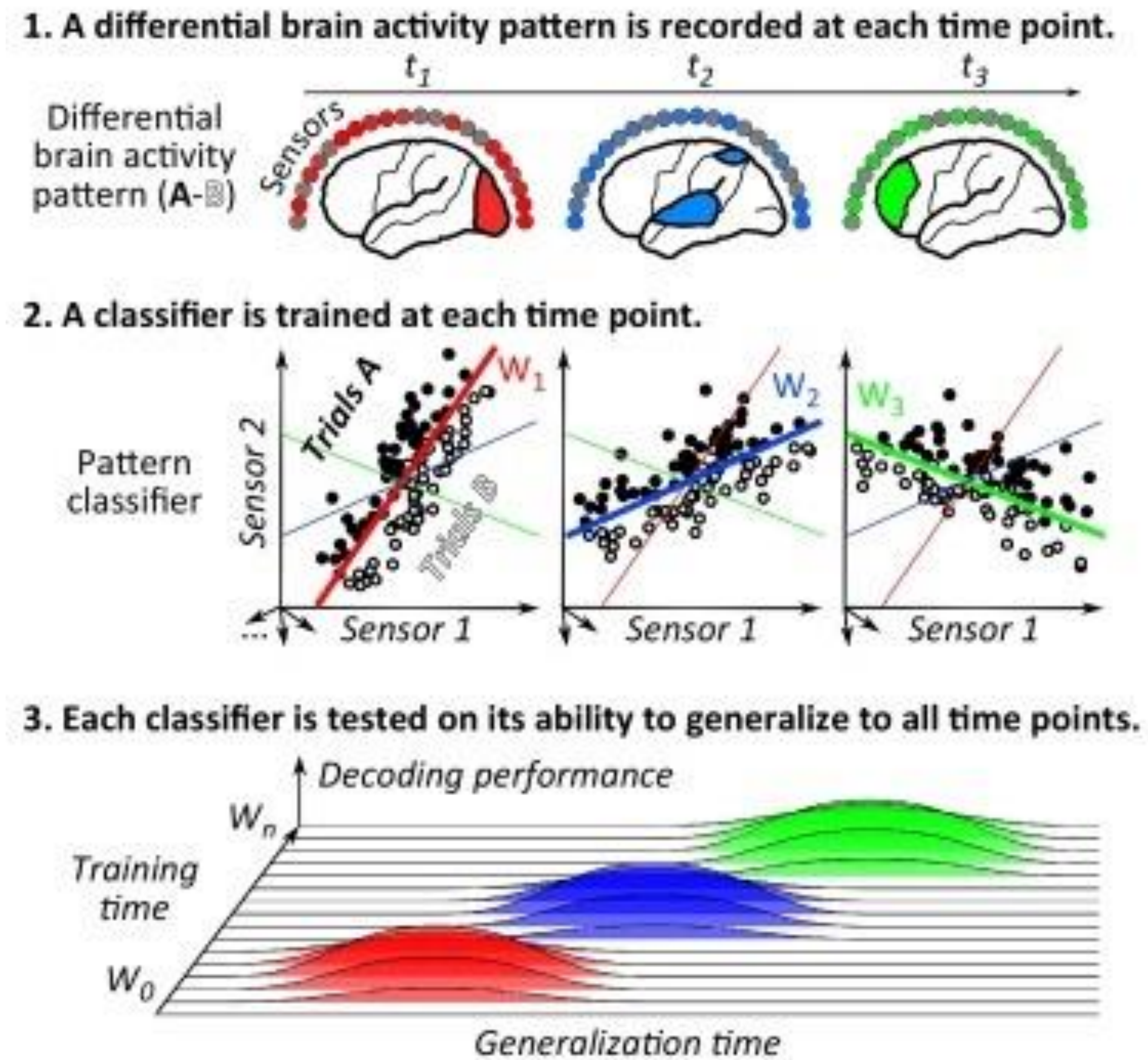


Figure 4. Illustration of the Multivariate Pattern Analysis (MVPA) analysis pipeline (reproduced with permission from King & Dehaene, 2014; see Appendix F).

For each time point, data from all sensors is aggregated and plotted in multidimensional space, with each dimension corresponding to a different sensor. Then, a classifier is trained to identify the boundaries that best divide stimuli from different classes of data. Classifier performance can subsequently be assessed not only by testing on the corresponding trial time point in the test data, but also by testing on previous and/or subsequent time points. This allowing for generalization of neural activation patterns across time.

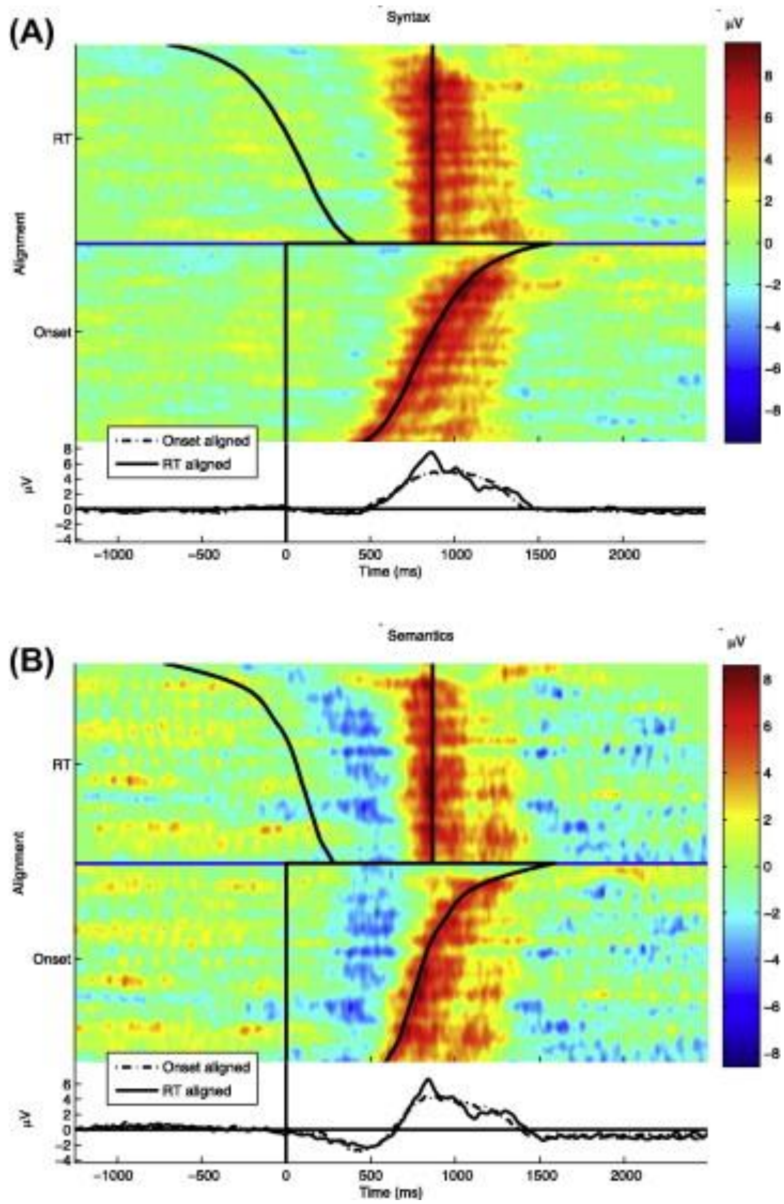


Figure 5. ERP images of difference waves (violation minus control difference) illustrating components that are vs. are not response-locked (reproduced with permission from Sassenhagen et al., 2014; see Appendix F). Note that Gaussian visual smoothing has been applied for presentation purposes.

Panel A shows a P600 component (in red), and Panel B shows an N400 component (in blue). For each panel, the top graph shows the data plotted in a response-locked fashion, i.e., so that 0 ms represents the behavioral response. Meanwhile, the bottom graph shows the data plotted in a stimulus-locked fashion, i.e., so that 0 ms represents onset of the stimulus.

As can be seen in Panel A, variability in P600 onset can be seen when data is represented in a stimulus-locked way, but this variability disappears when data is represented in a response locked-way. By contrast, in Panel B variability in N400 onset is seen in the response-locked plot but not the stimulus-locked plot.

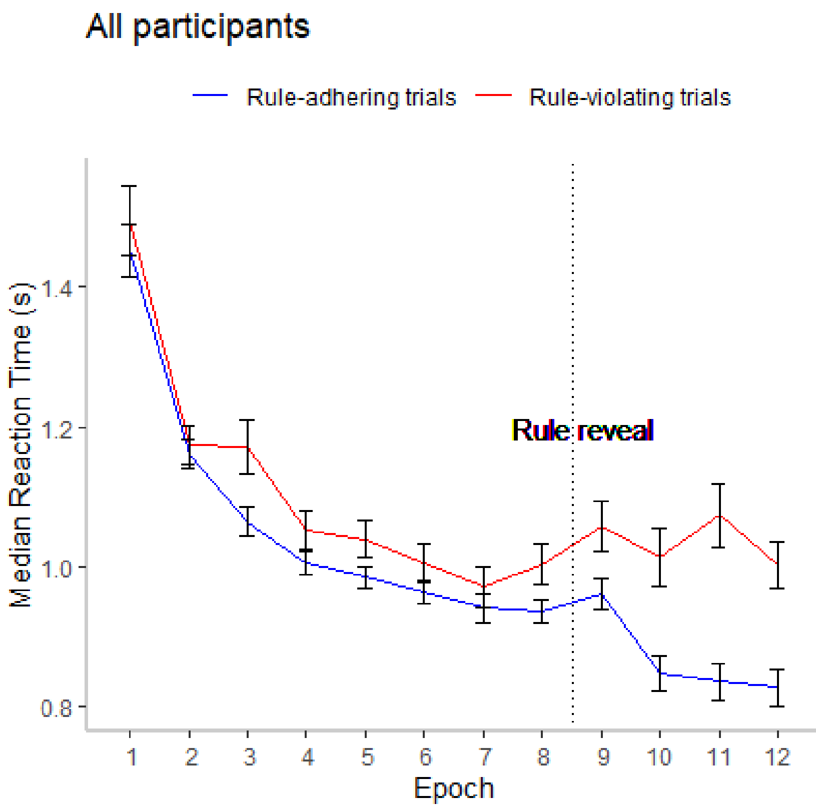


Figure 6. Participants' median response times to rule-adhering (rule-conforming) vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

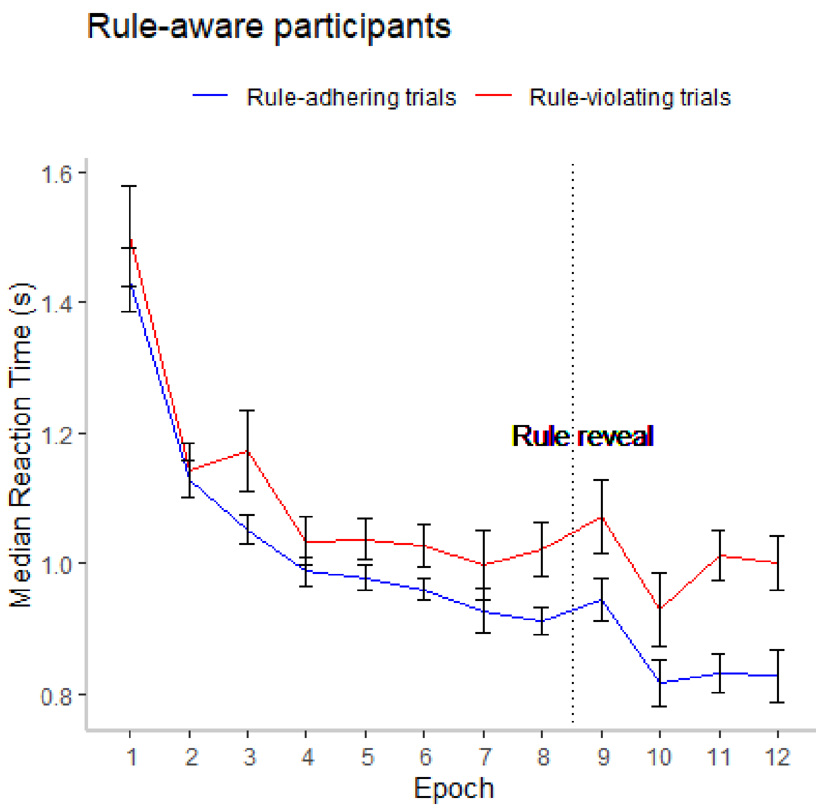


Figure 7. Rule-aware participants' median response times to rule-adhering vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

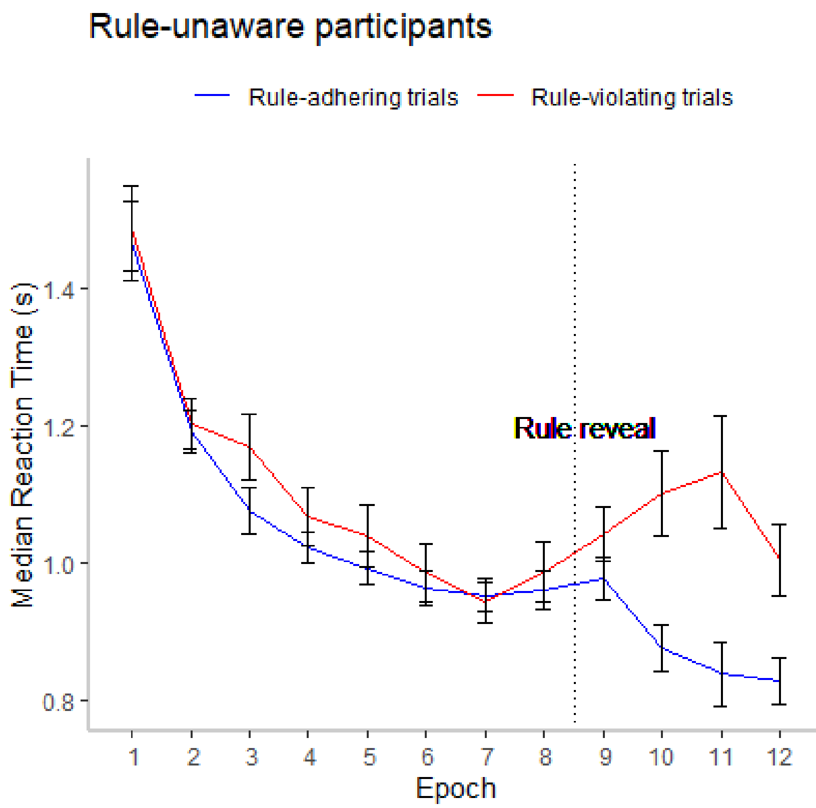


Figure 8. Rule-unaware participants' median response times to rule-adhering vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

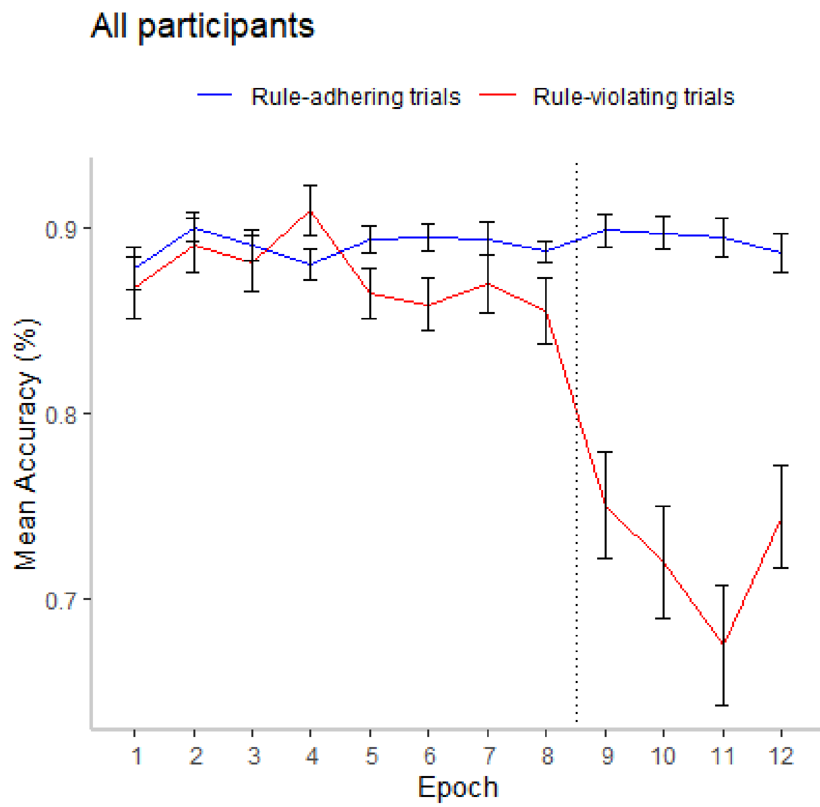


Figure 9. Participants' mean accuracies to rule-adhering (rule-conforming) vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

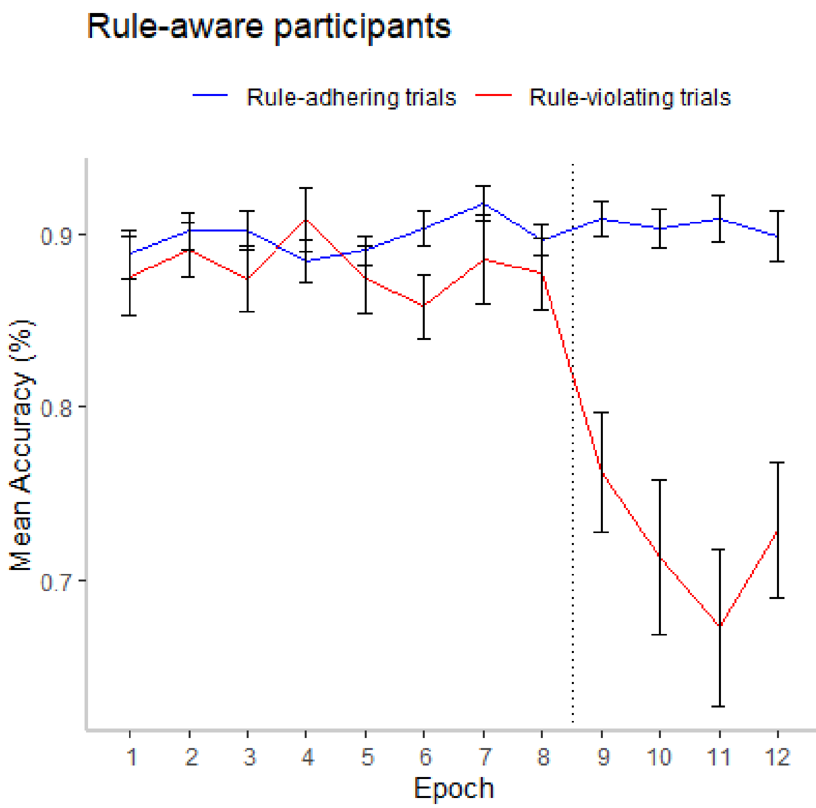


Figure 10. Rule-aware participants' mean accuracies to rule-adhering vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

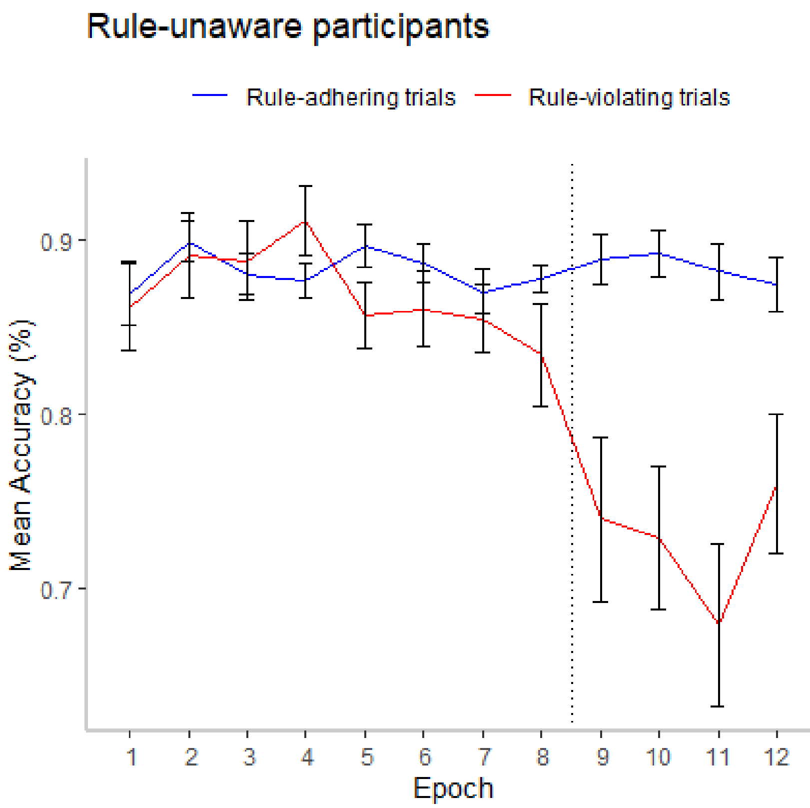


Figure 11. Rule-unaware participants' mean accuracies to rule-adhering vs. rule-violating trials, calculated by epoch. Error bars indicate standard errors.

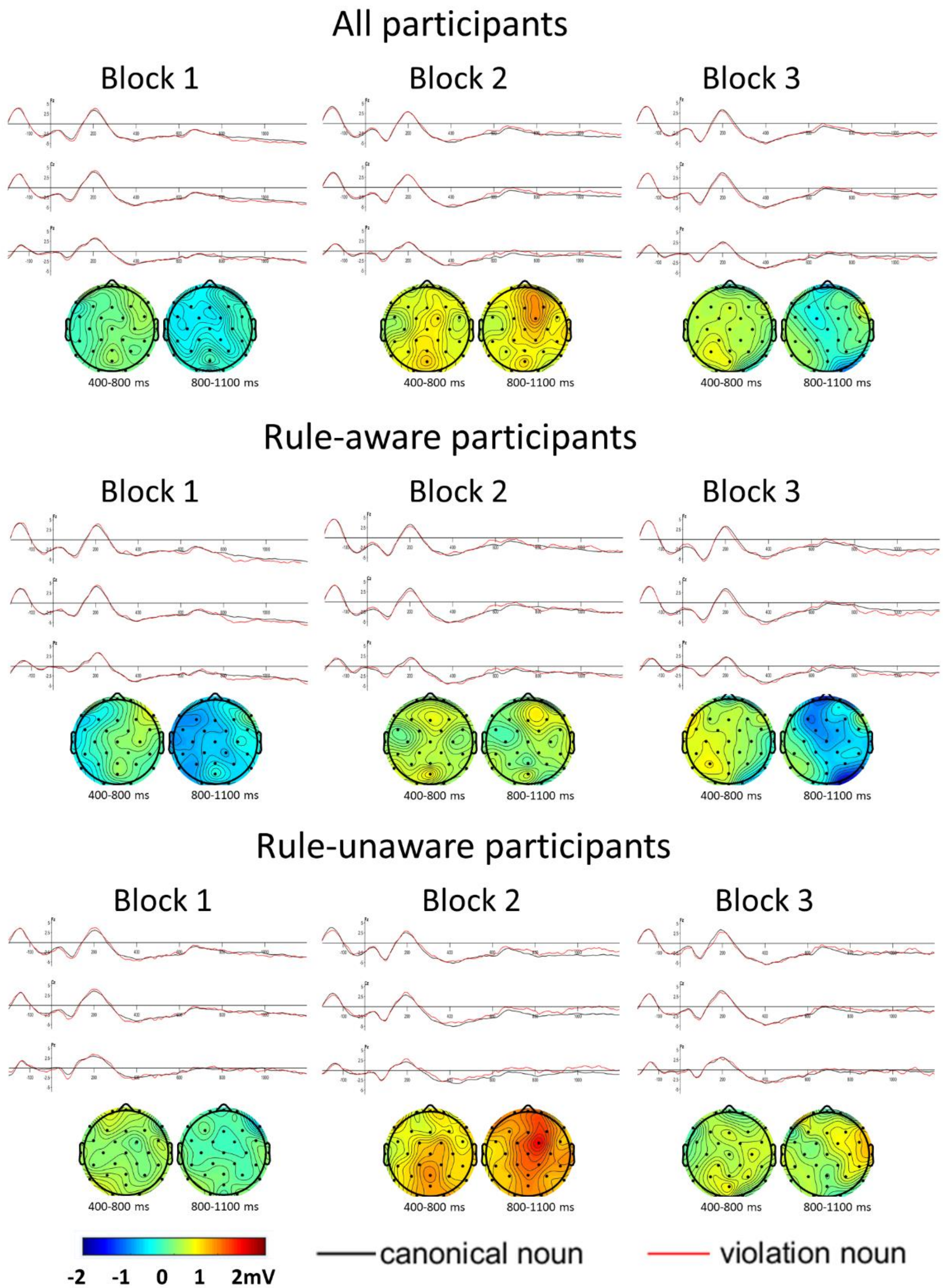


Figure 12. ERP waveforms and scalp maps comparing rule-adhering vs. rule-violating trials.

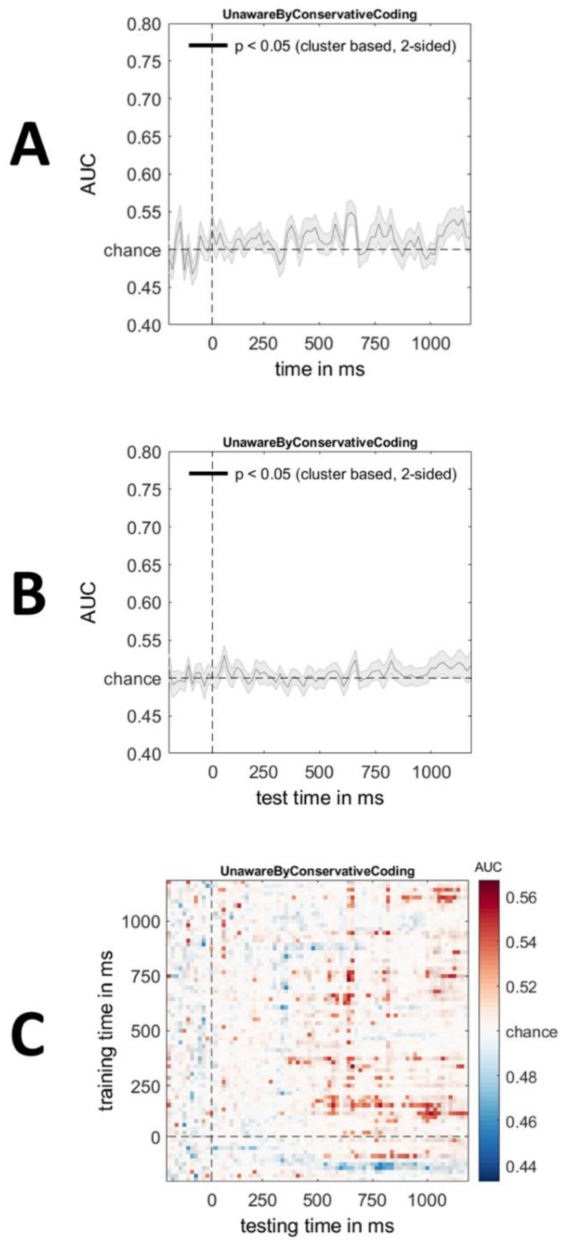


Figure 13. Trial classification performance of an MVPA decoder trained on rule-unaware participants' Block 2 data (prior to the debriefing interview) and tested on those same participants' Block 3 data (after they were told the hidden grammar rule). Panel A shows performance in diagonal decoding, i.e., when training and testing are performed on the same time points without any generalization across time windows. Panel B shows performance when the decoder is trained solely on neural activity in the time window 800-1100 ms (in which the late positivity was detected in my midline analysis and in Batterink et al., 2014). Panel C shows the temporal generalization matrix, which indicates how well neural activity for each time point in the training data generalizes to any given time point in the testing data. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold. However, as no effects were significant, nothing is highlighted.

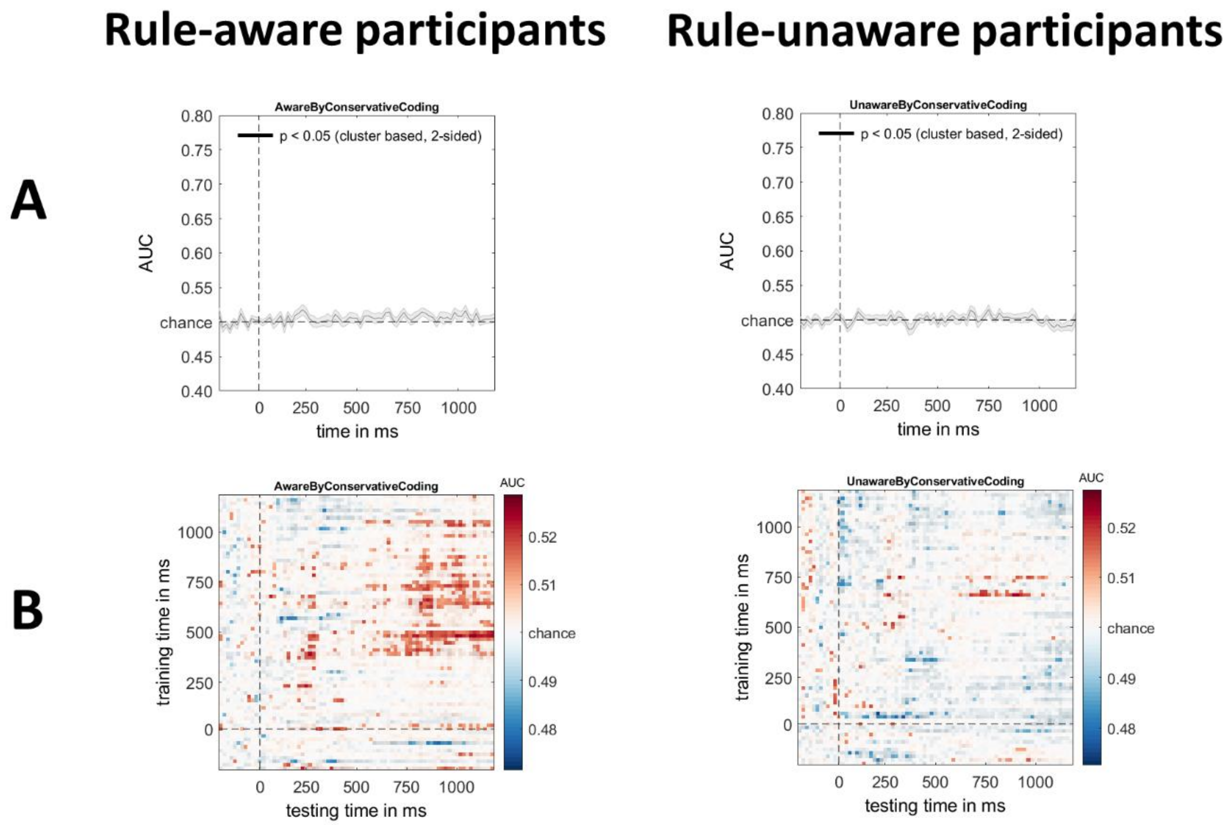


Figure 14. Trial classification performance of an MVPA decoder trained on the animate/ inanimate distinction in the noun-only block and tested on the artificial language articles in Block 2, shown separately for rule-aware and rule-unaware participants. The plots in Panel A shows performance in diagonal decoding, i.e., when training and testing are performed on the same time points without any generalization across time windows. The plots in Panel B show temporal generalization matrices, which indicate how well neural activity for each time point in the training data generalizes to any given time point in the testing data. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold. However, as no effects were significant, nothing is highlighted.

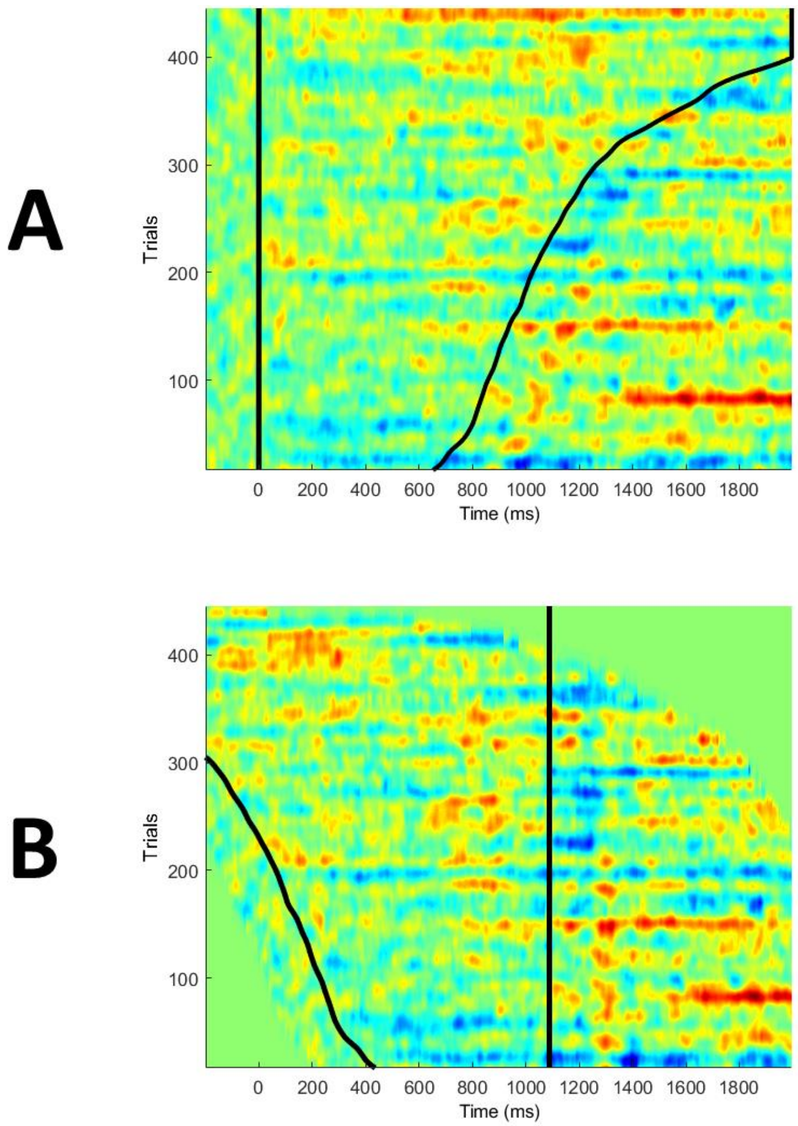


Figure 15. ERP images showing, for rule-unaware participants' Block 2 data, each rule-violating trial's difference wave relative to the average of the rule-adhering trials for that participant. Trials are sorted so that trials with faster response times are on the bottom and trials with slower response times are on top. Panel A shows the data plotted in a stimulus-locked fashion, so that the vertical black line at 0 ms represents the onset of word presentation and the curved black line on the right represents the participant's reaction time for that trial. Note that the line representing reaction times becomes vertical in the top-right of the plot because some reaction times were above 2000 ms and thus could not fit in the figure. Panel B shows the same data "shifted" in a response-locked fashion, so that the curved black line on the left represents the onset of word presentation and the vertical bar represents the reaction time for that trial.

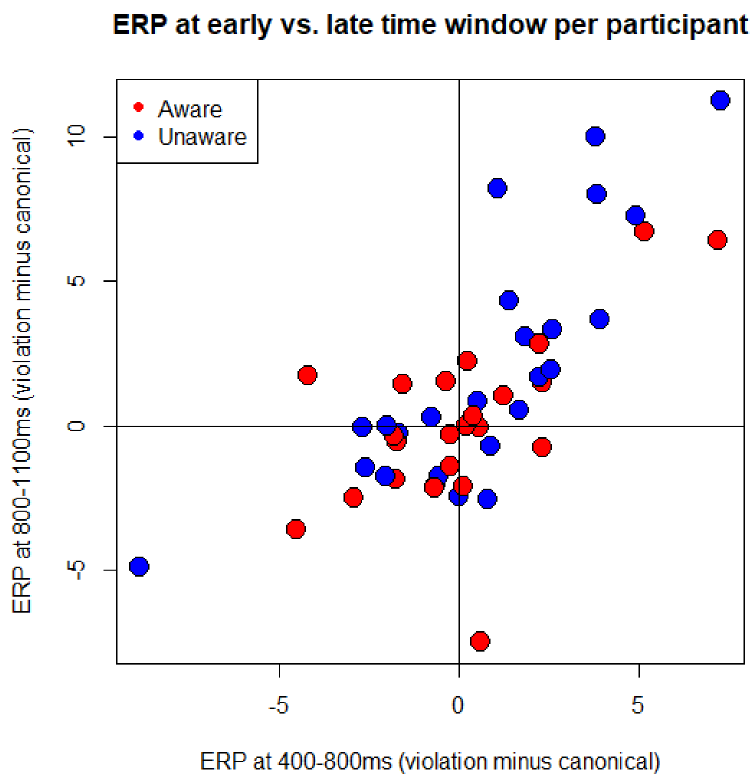
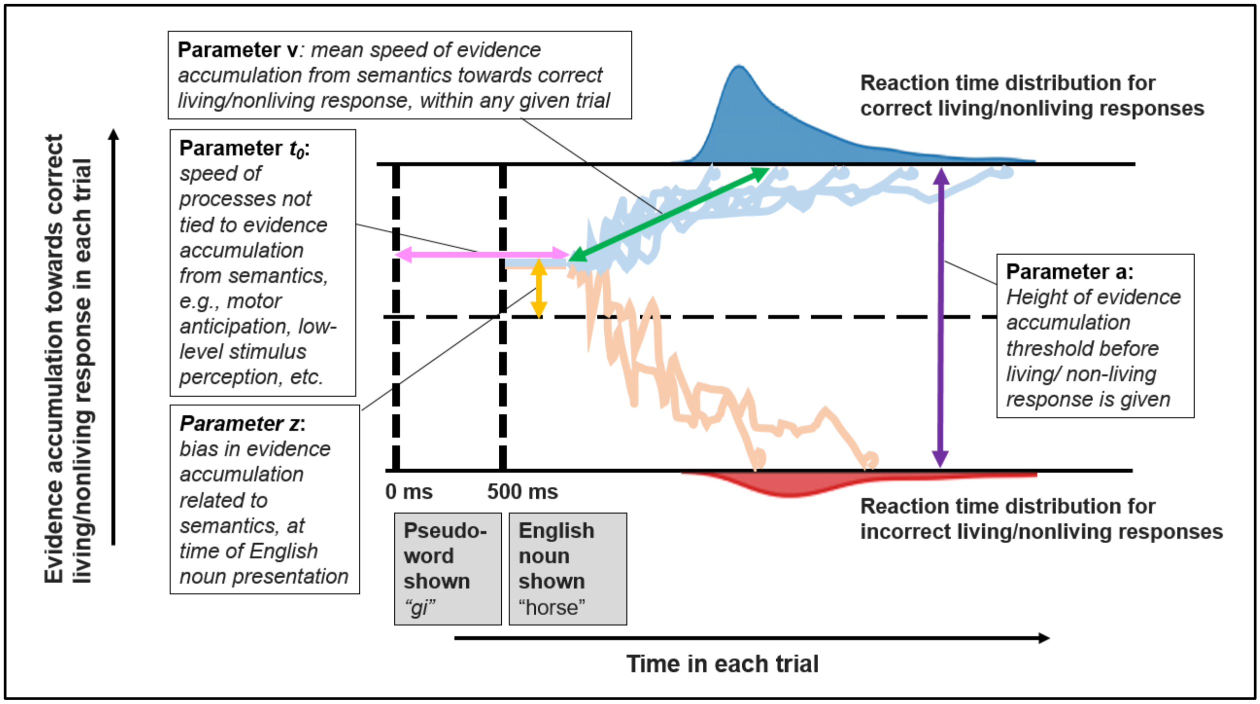


Figure 16. Panel A: ERP effects at early vs. late time windows per participant at electrode Cz in my data.

A



B

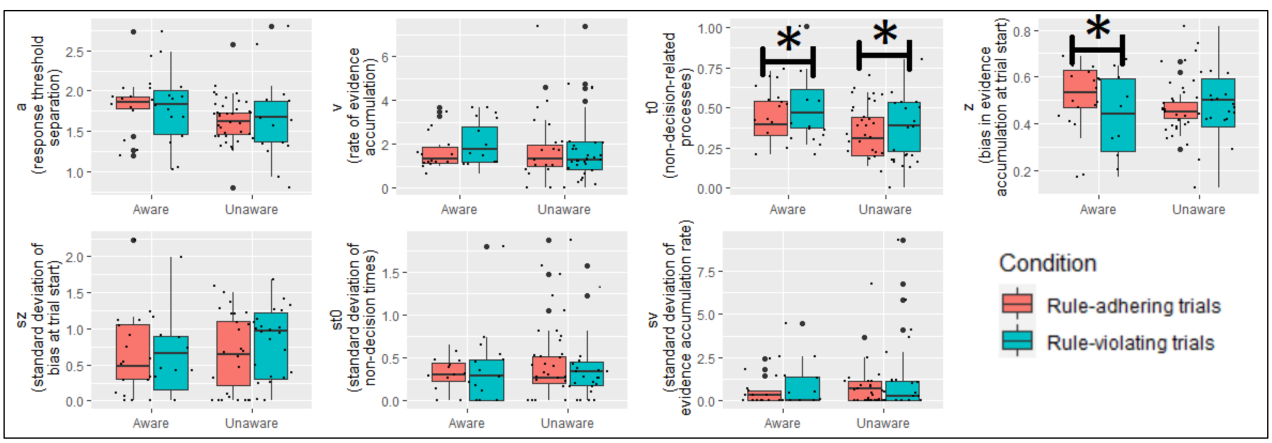


Figure 17. Panel A: Visualization of drift-diffusion model in the context of my experiment paradigm. Not pictured: parameters capturing standard deviation of v, t_0, and z.

Panel B. Drift-diffusion model parameter estimates for rule-adhering vs. rule-violating trials, shown separately for rule-aware vs. rule-unaware participants.

Reproduced from Abugaber and Morgan-Short (2021).

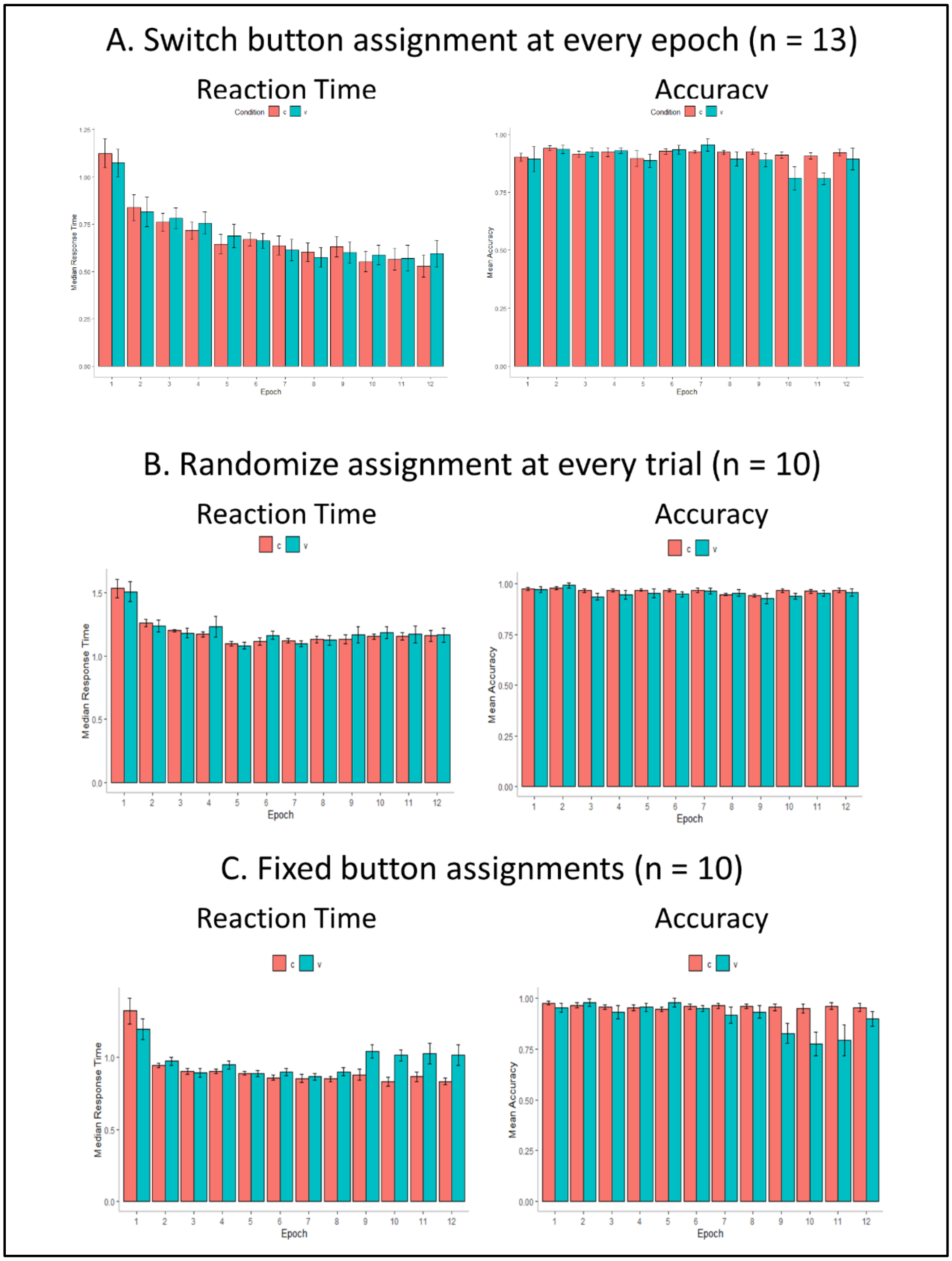


Figure 18. Epoch median reaction times and epoch mean accuracies for three different versions of the experiment conducted to determine whether the experiment paradigm could be adapted to avoid a learning effect based on button-pressing contingencies. Panel A: version of the experiment wherein the living/nonliving button assignment is switched at every epoch. Panel B: version of the experiment wherein the living/nonliving button assignment is randomized for every trial. Panel C: version of the experiment wherein the living/nonliving button assignment is fixed (as in the EEG experiment reported in this dissertation. Consistent behavioral effects were only found for the fixed-button version.

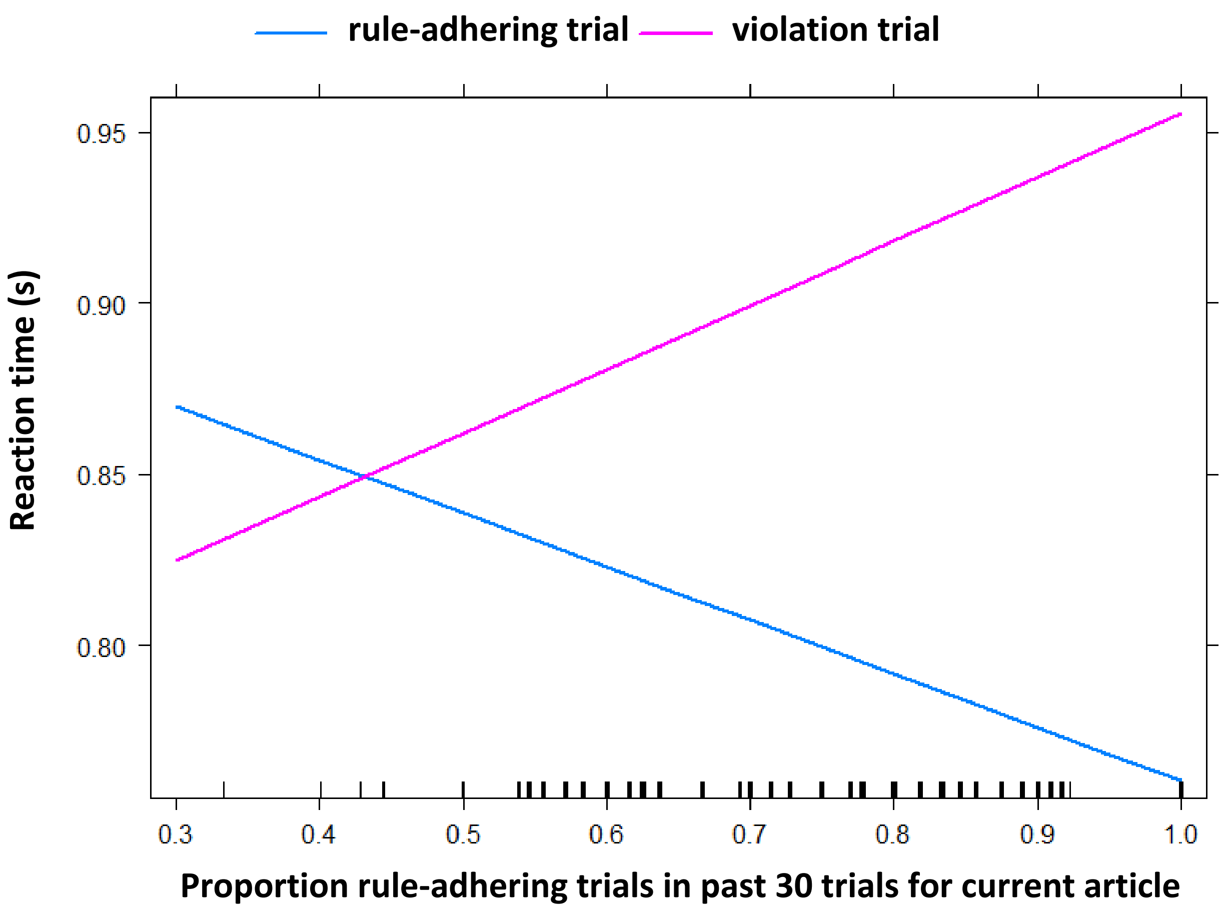


Figure 19. Visualization of how, in my reaction time data, trial condition (rule-adhering vs. rule-violating) interacts with the proportion of rule-adhering vs. rule-violating trials that a participant had seen for a pseudoword in the preceding thirty trials. As the proportion of rule-adhering trials goes up, reaction times for rule-violating trials (shown in pink) are slower and RTs for rule-adhering trials (shown in blue) are faster.

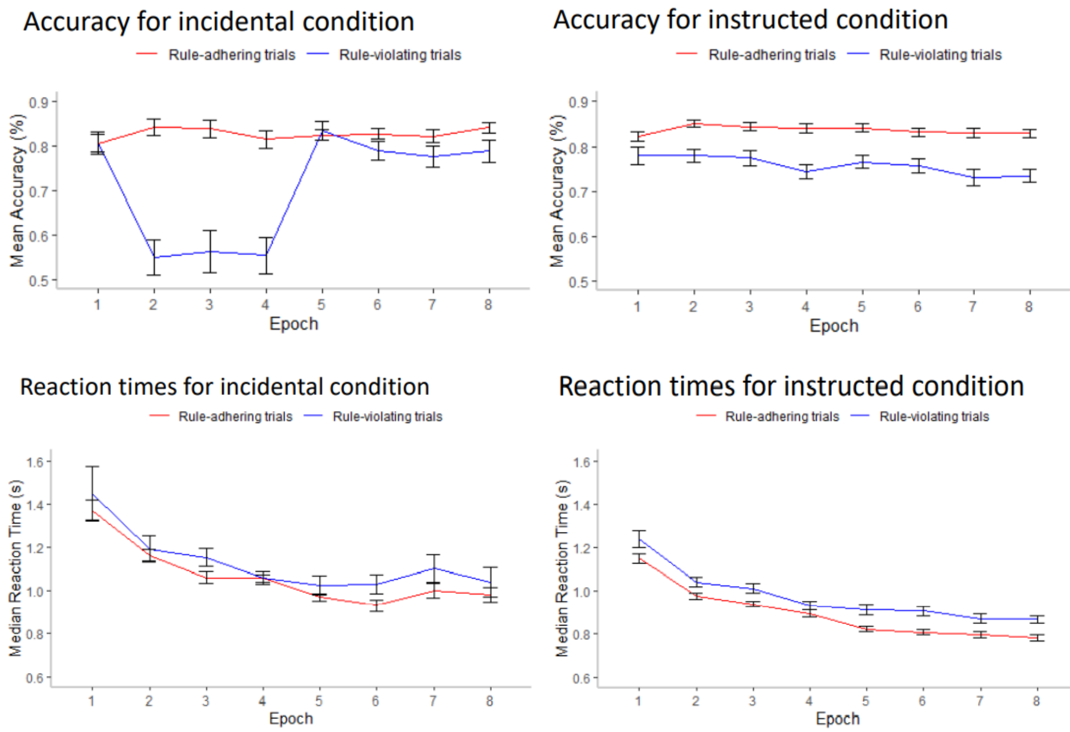


Figure 20. Median reaction times and mean accuracies for participants who did (instructed condition) vs. did not (incidental condition) receive explicit instruction about the hidden regularity in our experiment. Note that a five-minute break was given between epochs 4 and 5, potentially explaining the sudden change in accuracy for the incidental condition as participants shifted their strategies.

Appendix A: Language background questionnaire

Section A: General Information

- Gender
- Date of birth
- Age
- Number of years of formal education
- Program of study/major
- History of concussion/head trauma
- History of attention deficit disorder (ADHD) diagnosis
- Whether participant regularly takes psychoactive medication
- Diagnosis with learning disorder, cognitive disorder, or other language disorder
- Hearing or visual impairments
- Cultures with which one identifies
- Race
- Highest level of education completed (e.g., BA, 3 yrs. of college, etc.)
- Program of study/major (e.g., Economics)
- Cultures with which participant identifies
- Father's age
- Father's job
- Father's highest level of education
- Mother's age
- Mother's job
- Mother's highest level of education
- Date of immigration to USA (if applicable)
- Whether parents were born in USA or abroad
- Ethnicity

Section B: Language Background

- Participant lists all known languages in order of acquisition
- Participant lists all known languages in order of dominance
- Percent of time exposed to each language
- Percent of time that participant would choose to read in each language
- Percent of time that participant would choose to speak in each language
- Where were you exposed to each language? [ ] Home [ ] School [ ] Other: ____
- Age first exposed to each language
- Age began acquiring each language
- Age became fluent in each language
- Number of years and months living in a place where each language is spoken
- Number of years and months living with a family where each language is spoken
- Number of years and months working/studying in an environment where each language is spoken
- Self-rated speaking proficiency in each language (1-10)
- Self-rated listening proficiency in each language (1-10)
- Self-rated reading proficiency in each language (1-10)

- How much interacting with friends contributed to learning each language (1-10)
- How much interacting with family contributed to learning each language (1-10)
- How much reading contributed to learning each language (1-10)
- How much language tape/self-instruction contributed to learning each language (1-10)
- How much watching TV/videos contributed to learning each language (1-10)
- How much listening to radio/podcasts/music contributed to learning each language (1-10)
- Extent of exposure to each language when interacting with friends (1-10)
- Extent of exposure to each language when interacting with family (1-10)
- Extent of exposure to each language when reading (1-10)
- Extent of exposure to each language while in language-lab/self-instruction (1-10)
- Extent of exposure to each language when watching TV/videos (1-10)
- Extent of exposure to each language while listening to radio/podcasts/music (1-10)
- Extent of perceived foreign accent when speaking in each language (1-10)
- Extent to which others would identify one as non-native speaker when speaking in each language (1-10)

Appendix B: Experiment instructions

Noun-only block instructions

Task 1:

Use the j and k keys to indicate whether each word refers to something/someone that is living or non-living.

IMPORTANT: Don't respond until you get the cue that says "Respond now!"

The response button assignment will be randomly changed in each trial. Sometimes j means "living" and k means "non-living." Other times, j means "non-living" and k means "living." Follow the prompt to know which means which. This way, you can't prepare your response ahead of time :D

Also, the experiment won't proceed until you input the correct answer.

Press any of these two keys to begin the task.

Vocabulary pre-training instructions:

In some languages, the distance of an object being referred to is reflected in the grammar. In this experiment, 'gi' and 'ro' are used with objects that are 'near,' and 'ne' and 'ul' are used with objects that are 'far.' For example, the watch that you are wearing could be referred to as 'ro watch', which would mean something like 'the-near watch.' Thus, these words combine the English meaning of 'the' with the meanings of 'near' and 'far.' Press Enter to continue.

First you will complete two tasks to help you learn the meanings of these 4 new words:

gi (near)

ro (near)

ul (far)

ne (far)

Press Enter to continue.

Task 2:

You will be presented with an English word. Indicate which new word is a match, paying attention to both meaning and the color that the English word is written in. Say the word out loud while pressing the appropriate button. From left to right, the keys are J K L ; for each of the options on the screen. Press Enter to continue.

Task 3:

You will be now listen to the novel words, played one at a time. Indicate whether each word means 'near' or 'far' by pressing the appropriate button. Press U for near and I for far. Please note that color is now irrelevant. Press Enter to continue.

Instructions for practice and experimental blocks

Task 4:

You will now be exposed to a large number of examples of phrases that use these 4 novel words. On each trial, you will be asked to indicate:

1. Whether the object is living (J key) or non-living (K)

2. Whether the object is near (U) or far (I)

Please respond as quickly and accurately as possible, as we are measuring your reaction times. Also, the experiment will not proceed until you input the correct answer. Press any of these 4 keys to continue.

Appendix C: Word stimuli used in experiment

Practice Words (Animate)	**Practice Words (Inanimate)**
bat, sea lion, sea urchin	chandelier, chimes, bumper
Buffer Words (Animate)	**Buffer Words (Inanimate)**
aphid, millipede	sailboat, tube
Experimental Words (Animate)	**Experimental Words (Inanimate)**
crab, deputy, mongoose, custodian, golfer, trapeze, badger, bailiff, referee, dove, twin, clerk, stagehand, hypnotist, tycoon, aristocrat, trader, beetle, offspring, broker, tourist, communist, scout, barnacle, greyhound, seagull, biker, weatherman, caretaker, foreigner, rooster, rabbi, bartender, gymnast, jackrabbit, damsel, negotiator, skater, translator, dolphin, boxer, bodyguard, magician, customer, ballplayer, monarch, adviser, catfish, inventor, raven, intern, toad, pony, servant, starfish, spinster, flamingo, canine, aide, insect, clam, barber, coroner, announcer, newscaster, nominee, repairman, apostle, pioneer, penguin, hen, finch, jockey, preacher, goat, glutton, archduke, copilot, otter, ancestor, mackerel, calf, landlord, messenger, hawk, sentinel, matador, songwriter, pooch, diva, serpent, florist, housewife, tutor, teenager, editor, deejay, stylist, gull, raccoon, filmmaker, performer, geisha, hare, janitor, shrimp, sailor, bumblebee, historian, bookkeeper, wrestler, seller, hippie, swimmer, chauffeur, celebrity, citizen, outpatient, duckling, collie, smuggler, piglet, attendant, tailor, prey, tuna, employee, gladiator, newcomer, chancellor, slug, moose, firefly, gerbil, spectator, homeowner, giraffe, publicist, sophomore, catcher, prophet, butcher, warden, muskrat, termite, drifter, yak, constable, brute, plaintiff, instructor, ram, visitor, manicurist, oncologist, swordfish, quail, groundhog, possum, merchant, runaway, author, neighbor, iguana, shellfish, informer, sloth, adolescent, cod, partridge, bulldog, countess, frog, litigator, lecturer, bloodhound, masseuse, caveman, usher, poodle, mink, foreman, mole, engineer,	millennium, grindstone, ravioli, federation, refinery, saucer, steel, import, sphere, manure, tripod, crown, butter, pickle, lotion, denim, wheelchair, petroleum, riverbank, ventilator, platform, orbit, reunion, driveway, cathedral, lagoon, bandage, titanium, pasta, tomb, snack, hovercraft, riddle, tambourine, gazebo, container, patent, landscape, stereo, vent, paddle, muffler, spectacle, talisman, port, reminder, supper, diary, opera, whistle, screen, tether, nacho, gasoline, rosary, corduroy, kebab, scale, plaza, frame, trumpet, bunker, goulash, photograph, temple, morgue, curtain, risotto, knapsack, tanker, doorbell, vermouth, calculator, firewood, stimulus, label, recycling, segment, sunglasses, rubble, leotard, pulley, copper, polygraph, plutonium, stitch, domino, peninsula, trash, blueprint, pellet, junkyard, chain, ballot, exhaust, hammer, spool, sedative, seatbelt, bungalow, kayak, honeycomb, pallet, folklore, wallet, auditorium, harpoon, aftershave, minibus, ribbon, path, speck, spice, campfire, hairbrush, guitar, tailgate, amber, fertilizer, website, satire, gateway, portion, alley, menu, laboratory, radius, beverage, microphone, canoe, spur, porch, barrel, shed, corkscrew, backyard, scissors, mixer, election, enchilada, heirloom, collar, capsule, generator, sail, interior, lipstick, pawnshop, prom, ruler, mozzarella, moonshine, magnet, remote, greenhouse, elevator, wristwatch, frontier, spray, rubber, fundraiser, ransom, freeway, rooftop, migration, landfill, diagram, projector, sunscreen, possession, infirmary, industry, stench, locker, trombone, terrace, cellophane, pier, tuba, silicon, paperwork,

bootlegger, dropout, programmer, ferret, headmaster, soprano, flea, squirrel, geek, waitress, orangutan, rhinoceros, pug, drunkard, chemist, traveler, villain, butterfly, pirate, elephant, admiral, brunette, infant, spaniel, deserter, scavenger, mite, loon, grizzly, squid, veteran, hound, colleague, dentist, singer, shopper, crewman, trainee, lark, contestant, stewardess, slob, athlete, ant, comrade, starlet, toddler, locksmith, biologist, miner, refugee, dragonfly, vegetarian, freshman, policeman, nephew, barracuda, hitchhiker, puppeteer, youngster, busboy, boar, doorman, trooper, elk, roommate, aviator, albatross, oaf, pedestrian, schoolgirl, ballerina, grub, analyst, gorilla, groom, worker, porcupine, bluebird, descendant, immigrant, diplomat, predator, pupil, hog, salmon, traitor, lumberjack, primate, disciple, trucker, vulture, pest, astronaut, godfather, tick, salesman, ambassador, moth, bureaucrat, stallion, fawn, typist, walrus, newt, rodent, viper, recruit, beaver, accountant, passenger, shark, wanderer, milkman, vendor, husky, postman, tenant, falcon, duchess, designer, ox, acrobat, researcher, geologist, gazelle, hyena, mercenary, cricket, critic, carpenter, individual, seal, participant, mogul, chimpanzee, jellyfish, lifeguard, cub, samurai, peasant, supervisor, pope, beluga, carp, ostrich, dingo, pelican, diabetic, navigator, cockroach, donkey, shepherd, lamb, wolverine, nobleman, umpire, bronco, hacker, sparrow, oyster, camel, architect, nanny, nursemaid, bellhop, hunchback, treasurer, benefactor, beggar, eyewitness, heir, elder, worm, parent, seaman, farmer, mare, statesman, blackbird, tortoise, cleric, hobo, composer, scholar, banker, physician, executive, adult, koala, rancher, novelist, llama, bridesmaid, herring, cashier, swine, buyer, monk, chameleon, neurologist, heifer, hippo, handyman, inmate, drummer, steer, eagle, marine, shoemaker, teammate,	suburb, piccolo, seesaw, sickle, drought, packet, pill, prototype, culture, chariot, flask, trident, plot, shrine, lactose, marble, pastrami, string, sauerkraut, marker, implant, disguise, cocaine, loophole, toolbox, ridge, bistro, globe, courtyard, crust, luggage, sulfur, musket, tutu, razor, stream, monsoon, fragrance, graph, grease, timber, toaster, ration, telegram, goggles, shoreline, dormitory, lasso, portfolio, conch, portal, picnic, museum, bicycle, movement, franchise, snapshot, cauldron, lobby, gasket, windmill, bathtub, shingle, academy, embassy, skateboard, scaffold, debris, recipe, wrapper, exterior, pearl, cavern, toga, entree, ruby, ritual, symphony, handshake, milestone, image, microchip, moped, methane, plaster, gelatin, softball, blazer, runway, ukulele, incubator, fruitcake, helium, liquid, quadrant, winch, nightstand, strategy, pacemaker, scrape, palace, balloon, plastic, crucifix, education, nightgown, spinach, formula, treaty, tower, classroom, luncheon, arena, rickshaw, samba, flashback, salve, chisel, zipper, hydrogen, pedal, heliport, submarine, flavor, turquoise, repellent, saloon, bench, notebook, mound, corral, omelet, saliva, sewage, campsite, mortar, gimmick, gallon, mulch, crepe, printer, fortress, mayonnaise, sash, crystal, postcard, hardware, broom, locket, metal, urinal, strudel, shipment, nicotine, turpentine, parcel, smog, firearm, stamp, institute, plantation, watchtower, mannequin, fiddle, manhole, iodine, icing, email, porridge, couch, faucet, wharf, cassette, stagecoach, microwave, rotation, glucose, gauntlet, playground, empire, trail, cobalt, version, tray, sink, slipper, battery, vest, tavern, perfume, statistic, offense, outhouse, dumbbell, respirator, veil, instrument, stopwatch, university, subway, pagoda, frisbee, yearbook, eraser, mouthpiece, helmet, silk, dumpster, sticker, difficulty, scratch, chamber, jewelry, podium, haystack, plateau, salon, canyon,

condor, eel, vigilante, paramedic, goldfish, trout, vet, ape, operator, supporter, minion, investor, leech, lice, sidekick, reindeer, survivor, ranger, roman, blowfish, grocer, terrier, councilman, consumer, technician, conductor, crow, journalist, niece, turtle, tomcat, sardine, kicker, classmate, kangaroo, pilgrim, runner, mule, butler, swallow, widow, mammal, civilian, chaplain, sheepdog, pheasant, parakeet, socialist, snail, jester, redhead, caller, medic, narrator, sheep, tadpole, bystander, mockingbird, gopher, runt, stepfather, pitcher, fisherman, mechanic, nun, dancer, sportsman, antelope, proprietor, sentry, mosquito, zebra, consultant, pigeon, whale, goose, linebacker, shrew, crusader, goalie, jackal, caterer, mailman, server, therapist, baron, premed, developer, patrolman, baboon, hamster, patriot, parrot, cardinal, lobster, bachelor, mortician, stowaway, sculptor, groupie, octopus, shaman, canary, pastor, orphan, poet, bishop, vagrant, locust, tot, pharmacist, granddaddy, sultan, weasel, painter, deer, scientist, tyrant, woodpecker, owl, wildcat	pantry, ramp, treadmill, trinket, luxury, coaster, coolant, surplus, wrench, glacier, flounder, tunnel, headline, highway, yogurt, platter, triangle, briefcase, satin, valve, roast, brunch, homeland, snorkel, whisk, sidecar, mural, trunk, vessel, cabinet, penalty, steam, equator, nickel, puzzle, wrinkle, blindfold, mullet, quicksand, haiku, province, rust, thermostat, convention, candle, spittoon, altar, bubble, tonic, scalpel, scam, diaper, banjo, absinthe, latrine, topaz, airliner, lunchbox, tuxedo, receipt, fillet, package, taffy, hideaway, tricycle, hairpin, thread, lifeboat, plow, lavatory, territory, necessity, penicillin, crosswalk, saltwater, output, buckle, cooler, relic, agreement, venue, remedy, vehicle, mustard, attic, keyboard, clutter, graphite, structure, stocking, turban, dialect, laundry, plug, oval, haircut, radiator, stipend, aluminum, tweed, flurry, hostel, sunburn, irrigation, sedan, hydrant, nylon, theater, gadget, sauce, screenplay, revision, nickname, forceps, replica, labyrinth, stretcher

Appendix D: Awareness debriefing questionnaire

- Did you notice any specific patterns in the use of the new words you learned (other than the near/far rule that I taught you)? If so, what?
  - *If participant says yes*
    - Please describe the pattern that you noticed in the use of the new words.
    - When did you become aware that this pattern was relevant (i.e., before the break, after the break, or during this interview)?
    - When exactly did you first notice this pattern? Using your mouse, click on the slider on the screen to indicate at what point in the experiment you first started noticing this regularity.
    - Did being aware of this regularity in any way affect your performance in this experiment?
  - *Regardless of whether participant said they noticed a pattern:*
    - At any point, did you look for rules to figure out when to use the ul/gi/ro/ne words?
    - There is actually a rule that determines the correct word choice—i.e., when to say "gi" instead of "ro" or when to say "ul" instead of "ne." If you had to guess, what do you think that rule is?
    - Which of the new words are used more often to describe living things? And which words are used more often to describe objects?
    - There is a rule that determines the correct choice of ul/gi/ro/ne in the majority of cases. Namely, gi and ul are usually used for living things, and ro and ne are usually used for nonliving things. Did you ever consider this possibility?

Appendix E: Design decisions and follow-up analyses for MVPA

Design decisions for results presented here

For the MVPA analyses presented here, the decoder performance measure used is Area Under the Curve (AUC), a metric from signal detection theory which weighs classifier performance based on the model's degree of confidence (i.e., distance from the decision boundary) for each instance of trial classification (Wickens, 2002). In this way, low-confidence decisions would contribute less to the AUC metric than high-confidence decisions, as opposed to treating accuracy for all classifications equally as in the Correctly Classified Responses (CCR) metric, which merely calculates the raw percentage of trials that were classified accurately. Chance AUC performance is always 50%, in contrast to the systematic above- or below-chance accuracies that have been reported for simulated null datasets when using CCR as metric of MVPA decoding performance (Jamalabadi et al., 2018).

The MVPA approach in this dissertation uses a Backward Decoding Model (BDM), which predicts an experimental variable (e.g., trial condition) when given an observed pattern of brain activity. I chose this over a Forward Encoding Model (FEM), which predict patterns of brain activity when given an experimental variable. FEMs are more useful for continuous experimental variables (e.g., color or orientation of visually displayed stimuli) that have a direct relationship with neural activity, allowing one to predict activation patterns even for stimuli that had not been included in the training dataset (or even shown to participants in the original experiment). As this does not fit the categorical nature of the experimental variables in this study, I opt for a BDM (for further discussion of FEM, see Foster et al., 2016, 2017).

To control for inflated Type I (i.e., false positive) error rates from multiple comparisons (e.g., Bennett et al., 2009), I opt for a cluster-based comparison method when testing MVPA decoder performance. In this approach, time points of adjacent data are grouped into clusters (based on continuously significant t-tests) which are subsequently validated against a null distribution of cluster sizes generated by randomly permuting the observed AUC (Maris & Oostenveld, 2007). As such, the number of hypothesis-related tests is limited to the number of observed clusters, meaning that the potential for false positives is reduced. However, note that other Type I error-correction methods are also available, which may produce different results based on the duration and size of any EEG effects (e.g., False Discovery Rate correction; see Fahrenfort et al., 2018).

All MVPA analyses were performed using cross-class balancing, i.e., such that there are equal trial counts for rule-adhering and rule-violating trials when training the decoder. This eliminates bias stemming from the fact that we had a larger number of rule-adhering trials than of rule-violating trials. To achieve this, I implement between-class *oversampling* (duplicating underrepresented trials). This prevents us from unintentionally training the decoder to simply predict the most frequently occurring class. Balancing trial conditions in this way has been found to improve performance in linear models using area-under-the-curve measures (Xue & Hall, 2015).

The analyses for this dissertation use a linear rather than a non-linear classifier. The distinction between these is based on the shape of the decision boundary that is calculated when separating trials from different conditions: linear classifiers use a hyperplane whereas non-linear classifiers use non-planar boundaries that are more complex and thus more flexible (as illustrated in Figure 21). However, this flexibility makes non-linear classifiers susceptible to model overfitting (i.e., to tuning themselves to irrelevant noise or other idiosyncrasies in the training data) when sample sizes and effect sizes are low, as is commonly the case in EEG research

(Jamalabadi, 2017). As such, I opt for a Linear Discriminant Analysis (LDA) classifier, which uses a standard algorithm that has been shown to perform well in comparison to other decoding approaches (Grootswagers et al., 2017).

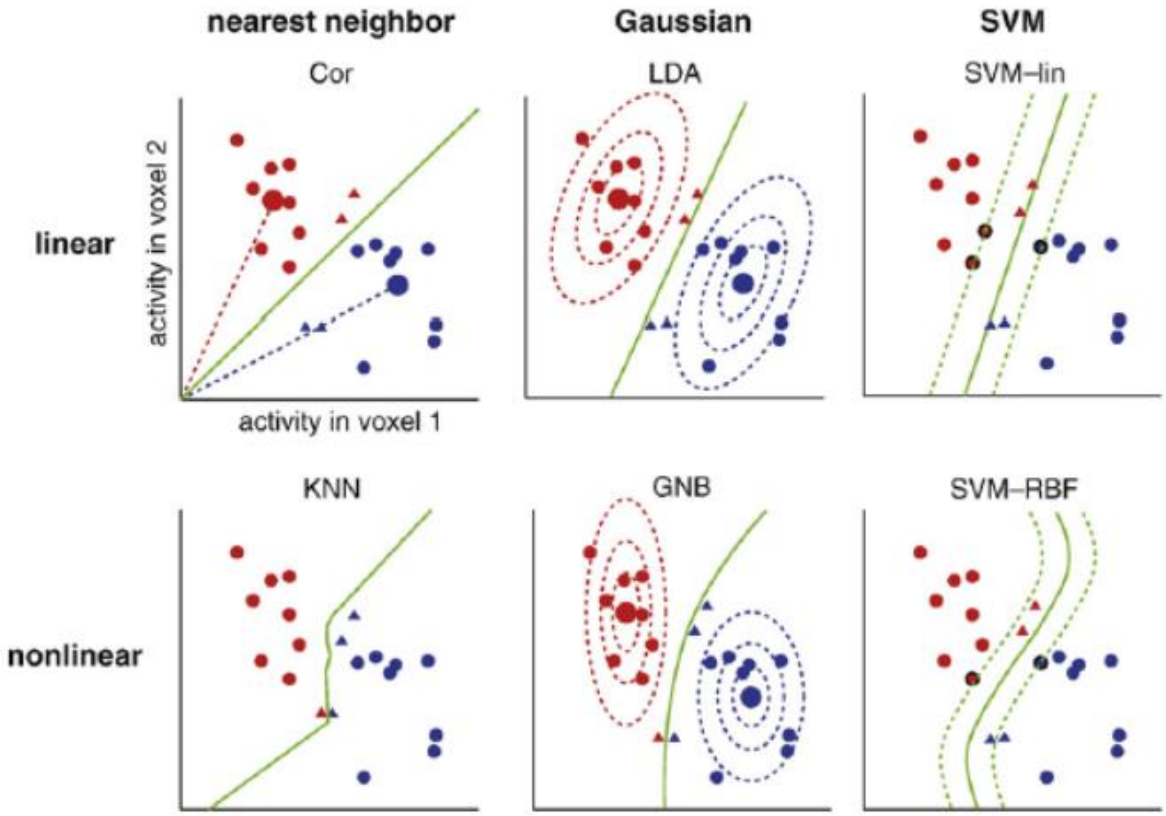


Figure 21. Illustration of linear classifiers (top row) and non-linear classifiers (bottom row) (reproduced with permission from Misaki et al., 2010; see Appendix F). The top row shows Correlation-based; Linear Discriminant Analysis; and linear Support Vector Machine classifiers. The bottom row shows *k*-Nearest Neighbor, Gaussian Naïve Bayes, and non-linear Support Vector Machine classifiers.

MVPA results in within-block decoding

The MVPA results reported in the dissertation are based on training an MVPA decoder on one section of the experiment (i.e., Block 2 for Research Question 2 and the noun-only block for Research Question 3) and determining whether this decoder achieves above-chance accuracy on a separate part of the experiment (i.e., Block 3 for Research Question 2 and the ul/gi/ro/ne pseudowords for Research Question 3). However, this analysis is predicated on the MVPA decoders being sensitive enough to detect differences in neural activity across experiment conditions (i.e., rule-violating vs. rule-adhering trials for Research Question 2 and animate vs. inanimate nouns for Research Question 3) in the first place, without attempting to generalize across different time points in the experiment.

To explore this, I performed MVPA classification with training and testing of the decoder on the same data instead of across data. To prevent overfitting of decoders to spurious noise, this within-block MVPA was validated using a *k*-fold cross-validation procedure (as illustrated in Figure 22 below). In this approach, part of the data is withheld at training and instead used to subsequently test the classifier. Several such "folds" (i.e., iterations) of training and testing are performed. For each iteration, different parts of the data are alternately used for training vs. testing, so that any particular trial is used for testing at least once over the course of the procedure. After running *k* such iterations, an average measure of decoder performance is calculated across these different iterations. This provides a measure of performance that is robust to idiosyncrasies in the training data (i.e., avoids "overfitting") and is thus more generalizable. I opt for five iterations of training/testing (i.e., a five-fold procedure), as this would strike a reasonable balance between achieving high trial classification accuracy vs. providing a stable, non-overfitted measure of classifier performance (Jamalabadi et al., 2016).

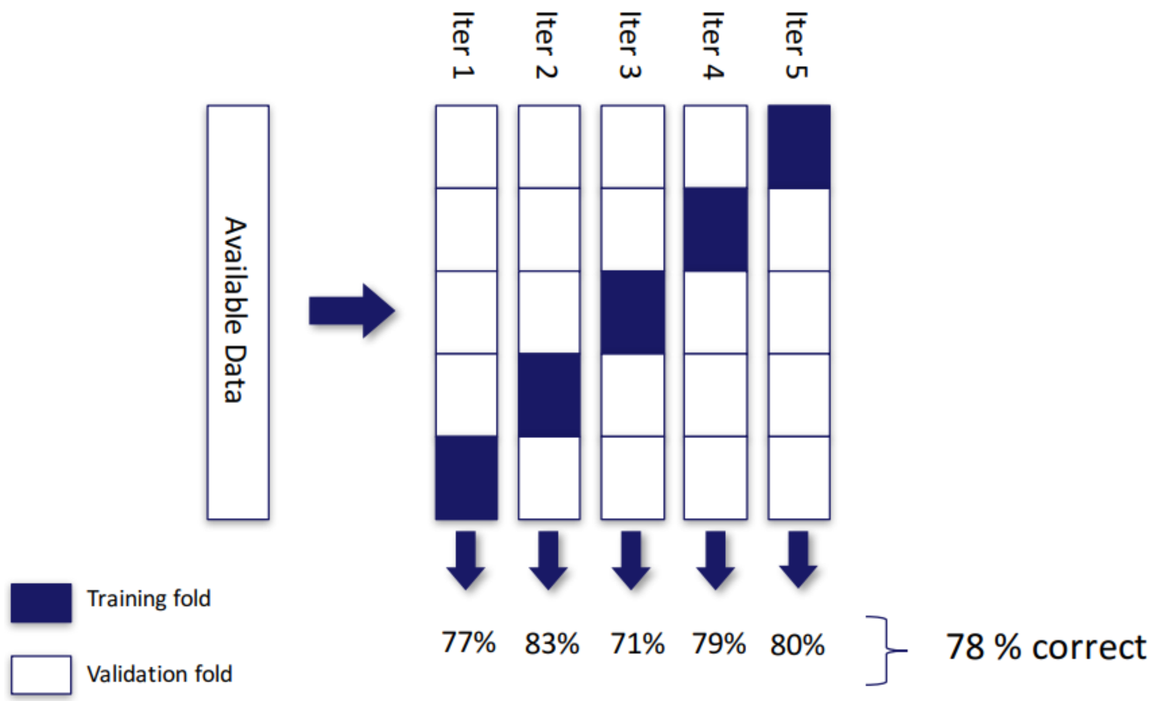


Figure 22. Illustration of *k*-fold cross-validation (in this case, five-fold validation; reproduced with permission from Hebart, 2016; see Appendix F). For each iteration of MVPA classifier training, one fifth of the data available for decoder training is withheld for testing. Decoder accuracy when testing on this left-out data is recorded before proceeding to the next iteration. An average of the classifier accuracies from each of these iterations of training and testing is then used as the final measure of classifier performance.

Unfortunately, when performing this within-block analysis, my MVPA decoders were not sensitive enough to pick up the relevant patterns in the data. Figure 23 shows the temporal generalization matrices for the noun-only block, Block 2, and Block 3, both for all participants as well as separately for rule-aware vs. rule-unaware participants. As can be seen, although the splotches of red in the heatmap suggests decoding accuracy above >50%, after cluster-based correction for multiple comparisons (see above) none of these decoders yielded significantly above-chance accuracy. This casts into doubt whether my results for Research Question 2 and Research Question 3 truly represent a lack of similarity between the training and testing data, or if the decoder is simply not sensitive enough to detect relevant neural activity in the training data to begin with.

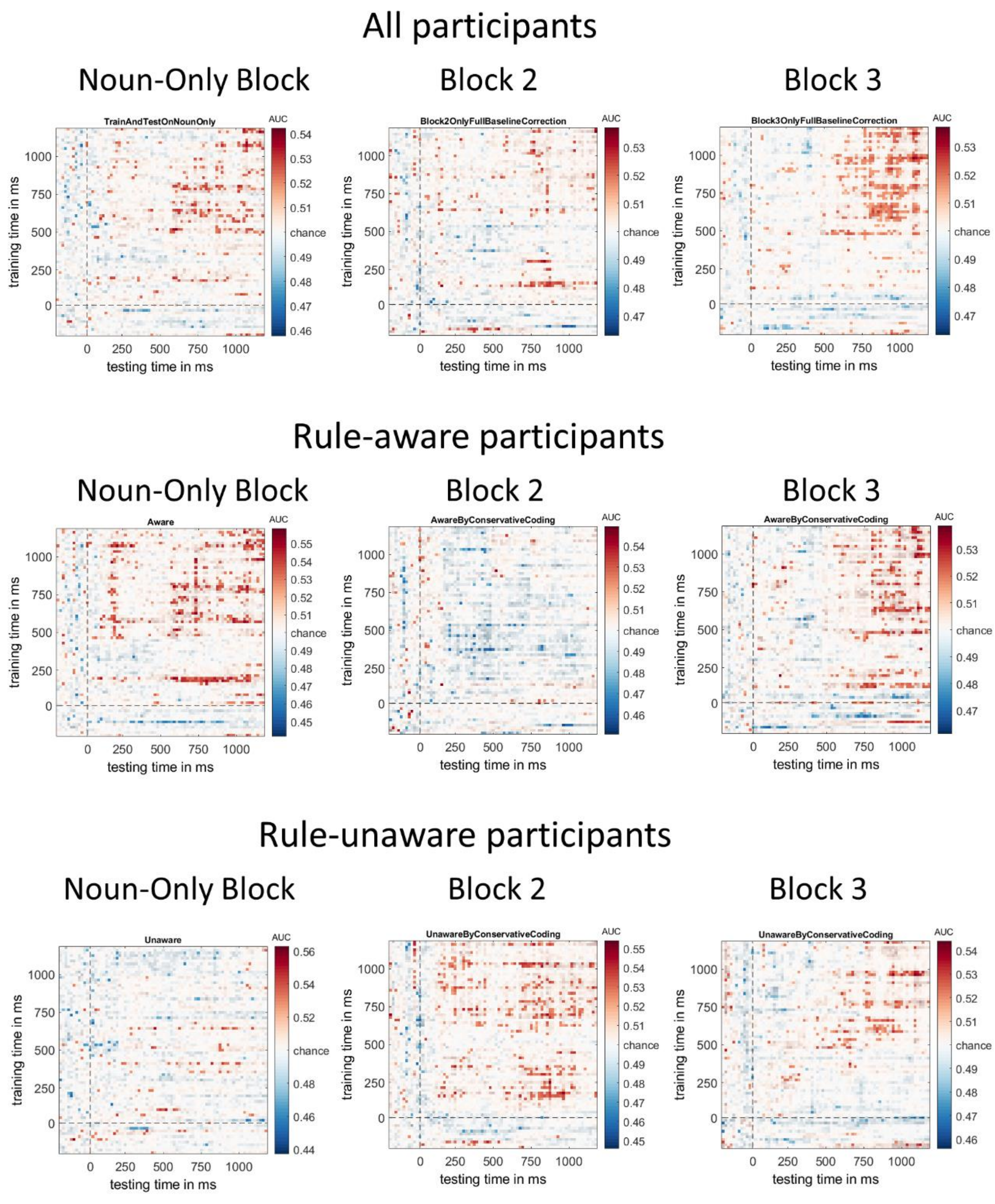


Figure 23. MVPA temporal generalization matrices for the animate/inanimate distinction in the noun-only block and for the rule-adhering/rule-violating trial distinction in Blocks 2 and 3. when training and testing the MVPA decoder on the same block of data. Shown for all participants as

well as separately for rule-aware and rule-unaware participants. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold. However, as no effects were significant, nothing is highlighted.

MVPA results with within-class balancing of ul/gi/ro/ne trials

The between-class balancing described above (i.e., keeping the number of rule-adhering and rule-violating trials equal when training my MVPA decoder) doesn't mean that subtypes are balanced *within* classes. To illustrate this: there is a possibility is that different proportions of *ul/gi/ro/ne* trials in the MVPA decoder training data led to one pseudoword being overrepresented in the decoder. In the worst-case scenario, this could lead to a situation where the decoder is essentially classifying trials as, e.g., "nouns that were preceded by *gi"* vs. "nouns that were not preceded by *gi"*, simply because the rule-violating and rule-adhering trials differed in their proportions of *gi* trials (see Görgen et al., 2018). To guard against this possibility, I present here the results of MVPA analyses that implement *within*-class balancing such that, within each class, the event subtypes were represented equally, i.e., so that I had the same number of *ul*, *gi*, *ro*, and *ne* trials in the rule-adhering as well as in the rule-violating data. This prevents the classifier from exploiting any subclass differences when distinguishing the two classes, e.g., by learning to classify *ul* trials as rule-adhering because *ul* trials were overrepresented in the rule-adhering trials used during classifier training. This within-class balancing was performed using within-class *undersampling* (throwing out overrepresented trials), which has the unfortunate consequence of reducing sample size (following the procedure performed by the ADAM MVPA toolbox; Fahrenfort et al., 2018). As seen below in Figure 24, I still did not yield statistically significant results when performing this within-class balancing. This means that a lack of statistical power from a small number of trials after balancing categories may have prevented us from detecting above-chance accuracies in neural decoding.

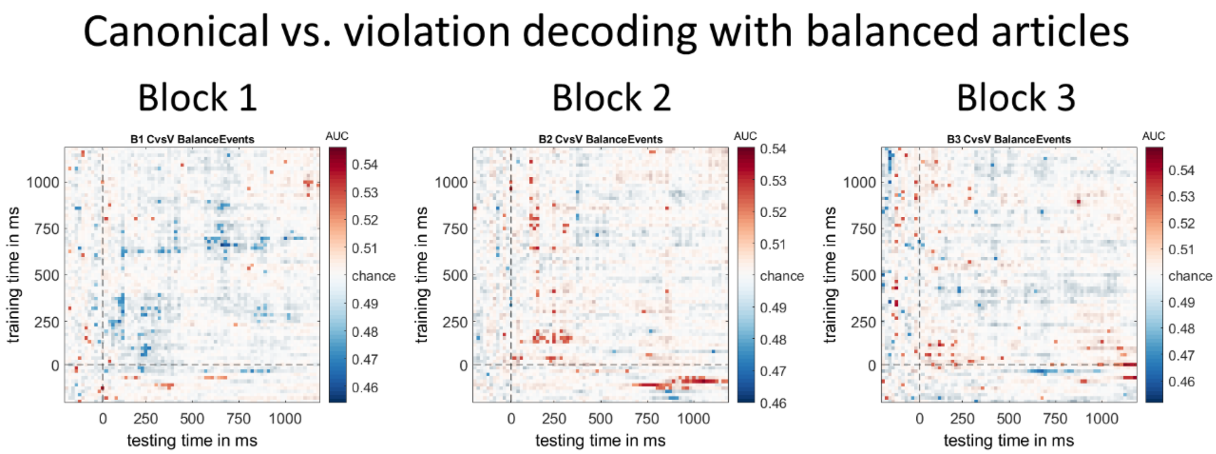


Figure 24. MVPA temporal generalization matrices for the rule-adhering/rule-violating trial distinction when the proportion of *ul/gi/ro/ne* trials in each condition is balanced. Shown separately for Blocks 1, 2, and 3. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold. However, as no effects were significant, nothing is highlighted.

MVPA separately on each of ul/gi/ro/ne

Another possibility is that participants paid specific attention to certain artificial language pseudowords more than to others. For instance, in my mid-experiment rule-awareness debriefing interview responses, many participants reported noticing a pattern for certain pseudowords more than others. Furthermore, one could conceptualize of the hidden rule not as a single cognitive entity that is learned by a participant, but rather as four separate contingencies (one for each of *ul/gi/ro/ne*). To account for this possibility that I had yielded null results because participants were simply learning the hidden rule for some articles but not others, I performed decoding in Block 2 of the data separately for each of the pseudowords *ul*, *gi*, *ro*, and *ne*. Although as seen in Figure 25 below the results based on raw decoding accuracy alone suggest that the rule was learned best for *gi*, moderately well for *ul* and *ne*, and not at all for *ro*, no results showed above-chance accuracy after cluster-based correction for multiple comparisons.

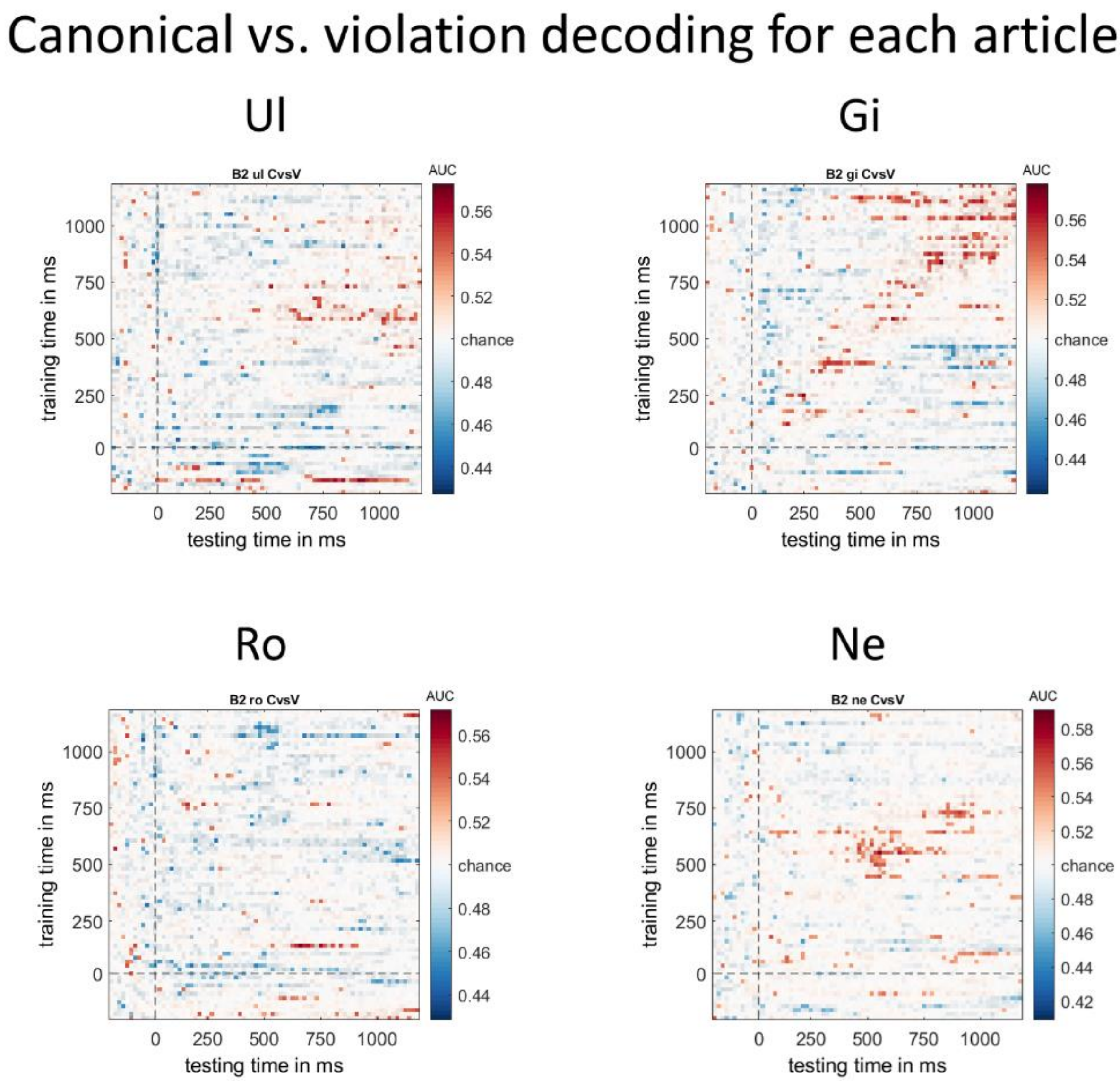


Figure 25. MVPA temporal generalization matrices for the rule-adhering/rule-violating trial distinction shown separately for each of *ul/gi/ro/ne* in Block 2 of the experiment. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold. However, as no effects were significant, nothing is highlighted.

MVPA results when using region-of-interest electrodes.

Another possibility is that neural activity related to the hidden grammar rule and related to the animate/inanimate status of nouns was present in my data, but only for specific electrodes. This is not implausible to the extent that both Batterink et al.'s (2014) as well as my results show that ERP components elicited by rule violations are stronger in some parts of the scalp than others. Thus, one way to improve my analysis (and mitigate the so-called "curse of dimensionality" from including so many different variables in my analysis) is to limit MVPA decoding to specific electrodes that are hypothesized to carry a particularly high level of decodable information in an exploratory analysis (Grootswagers et al., 2017). In the results presented below, I show no significant above-chance accuracy when using only electrodes previously identify to show language-related ERP effects (e.g., Tanner, 2019; Alday et al., 2017, Grey et al., 2017, Kim et al., 2018, Laszlo & Federmeier, 2014, Payne et al., 2015, Tanner, Inoue, et al., 2014, Tanner & Van Hell, 2014), namely: C3, Cz, C4, CP1, CP2, P3, Pz, and P4. As shown in Figure 26, I did not see signs of significantly above-chance trial classification with this approach.

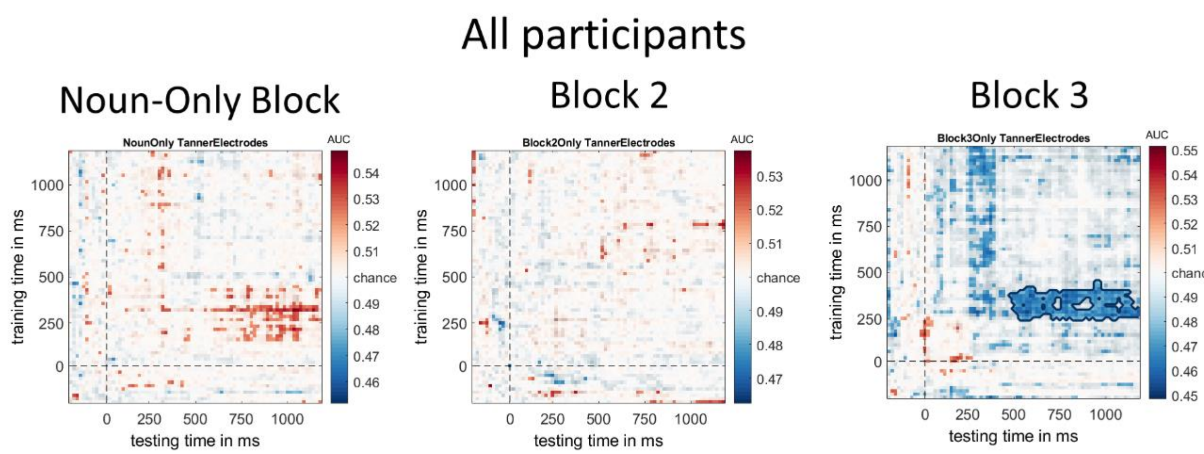


Figure 26. MVPA temporal generalization matrices for the animate/inanimate distinction in the noun-only block and for the rule-adhering/rule-violating trial distinction in Blocks 2 and 3 when using an electrode region of interest previously associated with language-processing effects (from Tanner, 2019: electrodes C3, Cz, C4, CP1, CP2, P3, Pz, and P4). Effects that are statistically significant after correction for multiple comparisons are highlighted in bold.

MVPA decoding on left vs. right button presses and for each button press assignment

One variable that is not of theoretical interest in my experiment is whether a given trial had a left vs. right button press. However, having an unbalanced number of left vs. right button press trials in my training data could be potentially problematic if these imbalances carried over to trial classification performance on my testing data. This is because, essentially, a decoder could inadvertently train itself to distinguish left vs. right trials (instead of, e.g., rule-violating or rule-adhering trials) because the training data did not have a proper balance of left vs. right trials.

To guard against this possibility, I performed *pre hoc* simulation analyses as per the "Same-Analysis Approach" (Görgen et al., 2018) wherein trials with a left button press had a mean amplitude of 5mV with random noise added, and trials with a right button press had a mean amplitude of 1mV with some random noise. In other words, this simulated the possibility that the only effects in my experiment were left/right button effects rather than rule-adhering /rule-violation trial effects. When I trained and tested a decoder on the rule-adhering/rule-violating distinction when using this simulation data, MVPA showed no significant above-chance accuracies either in diagonal decoding (Figure 27, Panel A) or in the temporal generalization matrix (Figure 27, Panel B). This confirms that the left/right button distinction would not lead to spurious effects in my data.

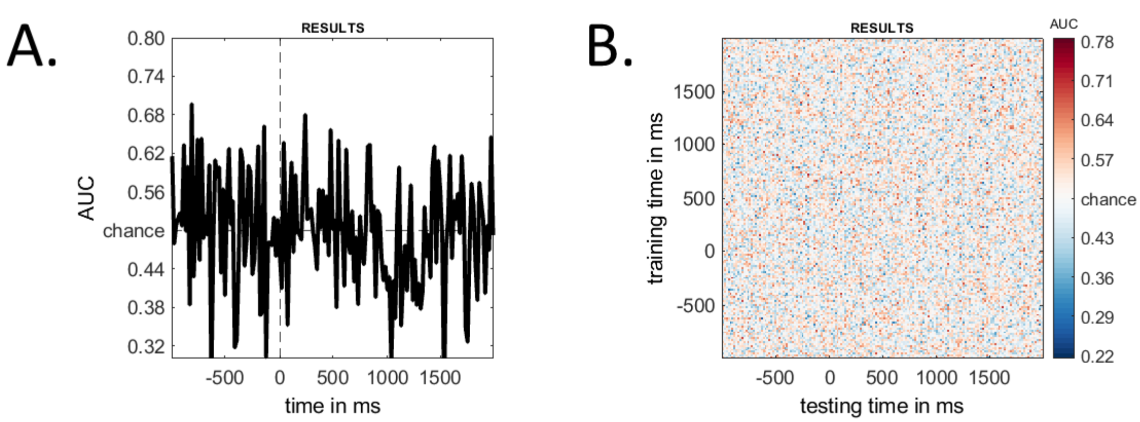


Figure 27. MVPA diagonal decoding performance (Panel A) and temporal generalization matrix (Panel B) for the rule-adhering/rule-violating trial distinction when simulating a strong effect only for the left or right button press in the experiment. Null results indicate no confound from the fixed button assignments in my experiment as far as decoding of rule-adhering vs. rule-violating trials is concerned.

As can be seen in the analyses on my real data shown below, the left/right button distinction could be decoded with above-chance accuracy in my actual data (see Figure 28, panel A). This is not surprising, and critically the simulation analysis described above confirms that this would not lead to spurious above-chance decoding accuracy in my results. That said, it is possible that variability between left/right button trials could overshadow variability between rule-adhering/rule-violation trials in my observed data, such that the rule-adhering/rule-violation MVPA analysis would yield null results. To guard against this possibility, I attempted to decode left-button trials that were rule-adhering vs. left-button trials that were rule-violating (Figure 28, panel B) as well as right-button trials that were rule-adhering vs. right-button trials that were rule-violating (Figure 28, panel C), under the intuition that within each of these analyses the data would be more homogenous because the same button press was involved. However, as can be seen below, none of these analyses yielded significantly above-chance results after cluster-based correction for multiple comparisons.

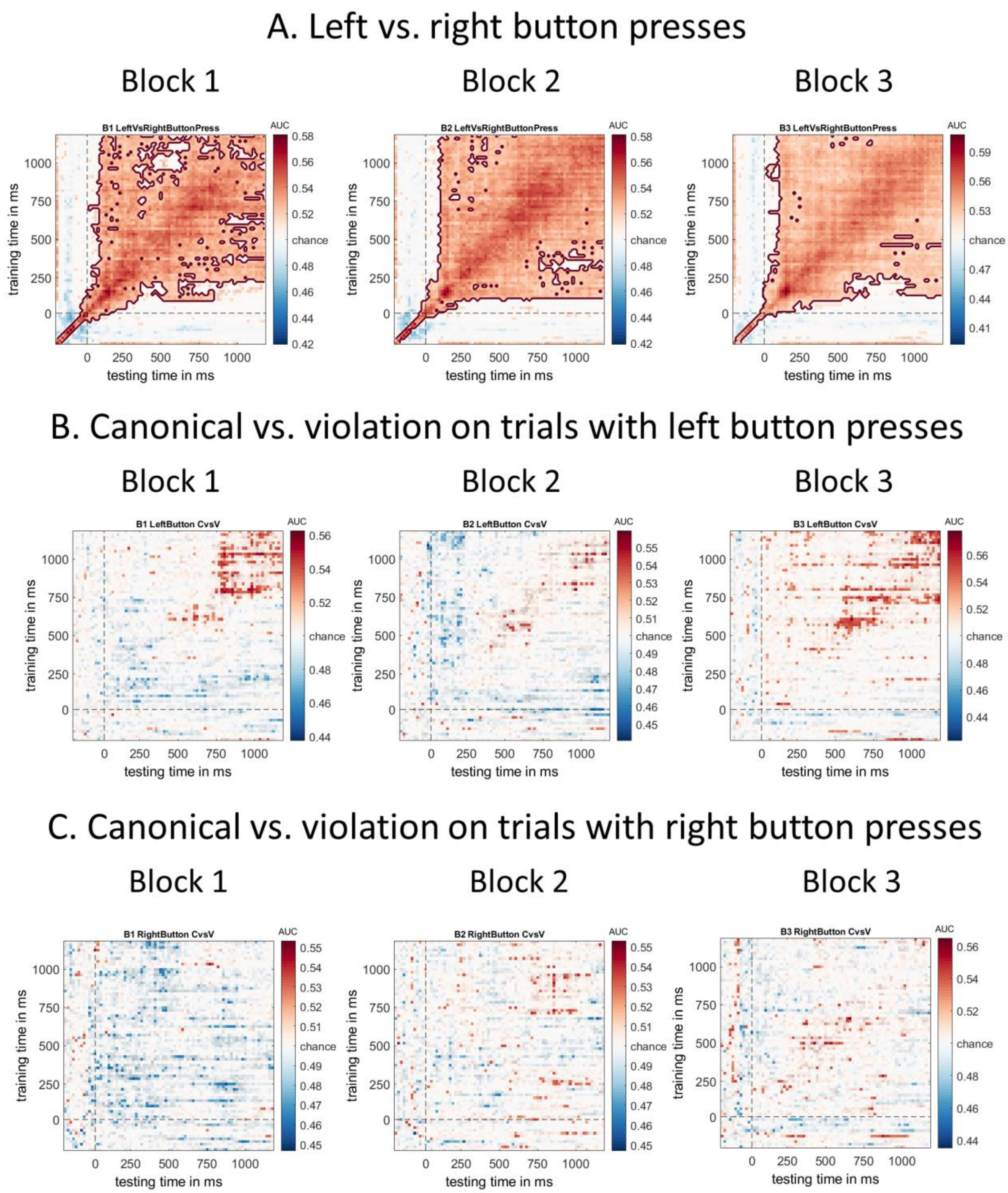


Figure 28. MVPA temporal generalization matrices for left vs. right button presses in Blocks 1, 2, and 3 of the experiment (Panel A). Rule-adhering vs. rule-violating trial decoding is shown separately for trials with a left button press (Panel B) and a right button press (Panel C). Effects that are statistically significant after correction for multiple comparisons are highlighted in bold.

MVPA on natural language grammar processing shows robust effects

As a “sanity check” to verify that the null results reported in this dissertation were not due to the specific parameters in my MVPA analysis, I performed the same decoding procedure on data from native speakers in an English sentence reading experiment (*N* = 52), with words shown one at a time as per a Rapid Serial Visual Presentation paradigm and a grammaticality judgment (“good/bad?”) after each trial. The analysis presented below in Figure 29 shows decoding results for grammatical vs. ungrammatical words, where ungrammaticalities were caused by a mismatch in determiner-noun number agreement (e.g., “this house” vs. “*this houses” or “these houses” vs. “*this houses.” As can be seen in the temporal generalization matrix in Figure 30 below, I found significantly above-chance trial classification accuracies starting at about 500 ms and proceeding for the rest of the EEG epoch up to a maximum of 1200 ms. This confirms that, all things being equal, my decoding approach is sufficient to detect neural activity associated with grammatical processing.

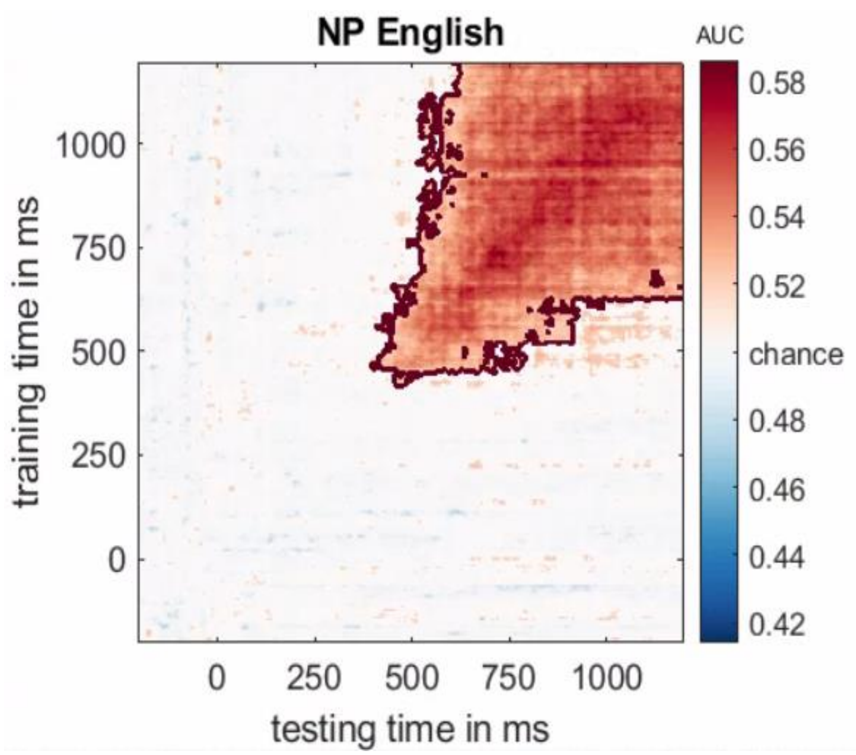


Figure 29. MVPA temporal generalization matrix for a grammatical vs. ungrammatical distinction on determiner-noun number agreement (e.g., *this house/these *house*) in an English sentence reading experiment. Effects that are statistically significant after correction for multiple comparisons are highlighted in bold.

Other possible future analyses

There are several future steps that could be taken in future MVPA analyses to attempt to improve decoder performance. First and foremost, Similarly, the time windows of decoder training can be "sharpened" into narrower bands than the broad 400-800 ms and 800-1100 ms time windows used here, e.g., by implementing Mass Univariate Analyses to determine the exact onsets and offsets of significant ERP differences and using only these narrower time windows for training my decoder (Groppe et al., 2011). Another way to address dimensionality issues is through a so-called temporal searchlight approach, i.e., by performing MVPA decoding on small clusters of consecutive time samples to increase the chances of detecting decodable patterns in the data (Su et al., 2012). In this way, one could achieve a principled balance between using individual time samples (which leads to the "curse of dimensionality") vs. generalizing training data from broad time windows (which throws out potentially relevant information from the exact of activity within this span). As a further possible refinement to my analysis approach, in the within-participants analysis presented here, five-fold cross-validation was performed. However, the number of folds could be adjusted, as this involves a trade-off between sensitivity to effects and overall decoder performance: previous simulations comparing the two possible extremes indicate that two-fold validation is more sensitive but yields lower overall accuracies than leave-one-out (LOO) approaches (which only withhold a single trial from each iteration of training) (Jamalabadi et al., 2016). As such, there are various other ways to approach the MVPA analysis besides the approach taken here.

Appendix F: Licenses for reprinting figures

CCC | Marketplace™

This is a License Agreement between David Abugaber ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Order General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.
All payments must be made in full to CCC in accordance with the Marketplace Order General Terms and Conditions below.

Order Date	08-Jul-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1245383-1	Publisher	ACADEMIC PRESS
ISSN	1095-9572	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	NeuroImage	Publication Type	e-Journal
Article Title	Comparison of multivariate classifiers and response normalizations for pattern-information fMRI.	Start Page	103
		End Page	118
Date	01/01/1993	Issue	1
Language	English	Volume	53
Country	United States of America	URL	http://www.journals.elsevier.com/neuroi...
Rightsholder	Elsevier Science & Technology Journals		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	Minor editing privileges?	No
Who will republish the content?	Academic institution	Incidental promotional use?	No
Duration of Use	Life of current edition	Currency	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Disentangling Neural Indices of Implicit versus Explicit Morphosyntax Processing in an Artificial Language	Institution name	University of Illinois - Chicago
		Expected presentation date	2022-07-22
Instructor name	David Abugaber		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	David Abugaber

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 1. Different pattern classifiers use different decision boundaries.	Title of the article/chapter the portion is from	Comparison of multivariate classifiers and response normalizations for pattern-information fMRI.
Editor of portion(s)	Misaki, Masaya; Kim, Youn; Bandettini, Peter A.; Kriegeskorte, Nikolaus	Author of portion(s)	Misaki, Masaya; Kim, Youn; Bandettini, Peter A.; Kriegeskorte, Nikolaus
Volume of serial or monograph	53	Issue, if republishing an article from a serial	1
Page or page range of portion	103-118	Publication date of portion	2010-10-14

CCC | Marketplace™

Order Date	08-Jul-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1245377-1	Publisher	Routledge
ISBN-13	9781134587537	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Explicit Learning in the L2 Classroom : A Student-Centered Approach	Country	United States of America
Author/Editor	Leow, Ronald P.	Rightsholder	Taylor & Francis Group LLC - Books
Date	03/20/2015	Publication Type	e-Book
Language	English	URL	http://www.tandfebooks.com/isbn/97813...

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
Format (select all that apply)	Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	No
Duration of Use	Life of current edition	Incidental promotional use?	No
Lifetime Unit Quantity	Up to 499	Currency	USD
Rights Requested	Main product		

NEW WORK DETAILS

Title	Disentangling Neural Indices of Implicit versus Explicit Morphosyntax Processing in an Artificial Language	Institution name	University of Illinois - Chicago
Instructor name	David Abugaber	Expected presentation date	2022-07-22

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	David Abugaber

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure "Model of the learning process in instructed SLA"	Title of the article/chapter the portion is from	N/A
Editor of portion(s)	N/A	Author of portion(s)	Leow, Ronald P.
Volume of serial or monograph	N/A	Issue, if republishing an article from a serial	N/A
Page or page range of portion	242	Publication date of portion	2015-03-19

CCC | Marketplace™

Order Date	08-Jul-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1245379-1	Publisher	Elsevier Science
ISSN	1879-307X	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Trends in cognitive sciences	Publication Type	e-Journal
Article Title	Characterizing the dynamics of mental representations: the temporal generalization method.	Start Page	203
		End Page	210
Date	01/01/1997	Issue	4
Language	English	Volume	18
Country	United Kingdom of Great Britain and Northern Ireland	URL	http://www.sciencedirect.com/science/jc
Rightsholder	Elsevier Science & Technology Journals		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	Minor editing privileges?	No
Who will republish the content?	Academic institution	Incidental promotional use?	No
Duration of Use	Life of current edition	Currency	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Disentangling Neural Indices of Implicit versus Explicit Morphosyntax Processing in an Artificial Language	Institution name	University of Illinois - Chicago
		Expected presentation date	2022-07-22
Instructor name	David Abugaber		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	David Abugaber

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure I. The principles underlying temporal decoding and temporal generalization	Title of the article/chapter the portion is from	Characterizing the dynamics of mental representations: the temporal generalization method.
Editor of portion(s)	King, J-R.; Dehaene, S.	Author of portion(s)	King, J-R.; Dehaene, S.
Volume of serial or monograph	18	Issue, if republishing an article from a serial	4
Page or page range of portion	203-210	Publication date of portion	2014-03-31

CCC | Marketplace™

Order Date	08-Jul-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1245378-1	Publisher	PUBLISHED BY THE MIT PRESS WITH THE COGNITIVE NEUROSCIENCE INSTITUTE
ISSN	1530-8898	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Journal of cognitive neuroscience	Publication Type	e-Journal
Article Title	Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data.	Start Page	677
		End Page	697
		Issue	4
Author/Editor	Cognitive Neuroscience Institute (Norwich, Vt.)	Volume	29
		URL	http://firstsearch.oclc.org/journal=0898-9...
Date	01/01/1989		
Language	English		
Country	United States of America		
Rightsholder	MIT Press - Journals		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	Minor editing privileges?	No
Who will republish the content?	Academic institution	Incidental promotional use?	No
Duration of Use	Life of current edition	Currency	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Disentangling Neural Indices of Implicit versus Explicit Morphosyntax Processing in an Artificial Language	Institution name	University of Illinois - Chicago
		Expected presentation date	2022-07-22
Instructor name	David Abugaber		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	David Abugaber

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 2. An illustration of how multivariate analysis can result in increased sensitivity compared to univariate analysis	Title of the article/chapter the portion is from	Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data.
Editor of portion(s)	Grootswagers, Tijl; Wardle, Susan G.; Carlson, Thomas A.	Author of portion(s)	Grootswagers, Tijl; Wardle, Susan G.; Carlson, Thomas A.
Volume of serial or monograph	29	Issue, if republishing an article from a serial	4
Page or page range of portion	677-697	Publication date of portion	2017-03-31

CCC | Marketplace™

Order Date	08-Jul-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1245381-1	Publisher	ACADEMIC PRESS
ISSN	0093-934X	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Brain and language	Rightsholder	Elsevier Science & Technology Journals
Article Title	The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned.	Publication Type	Journal
		Start Page	29
		End Page	39
Date	01/01/1974	Volume	137
Language	English, English		
Country	United States of America		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	Minor editing privileges?	No
Who will republish the content?	Academic institution	Incidental promotional use?	No
Duration of Use	Life of current edition	Currency	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Disentangling Neural Indices of Implicit versus Explicit Morphosyntax Processing in an Artificial Language	Institution name	University of Illinois - Chicago
		Expected presentation date	2022-07-22
Instructor name	David Abugaber		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	David Abugaber

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure. 2. RT-sorted ERP images of difference waves at PZ	Title of the article/chapter the portion is from	The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned.
Editor of portion(s)	Sassenhagen, Jona; Schlesewsky, Matthias; Bornkessel-Schlesewsky, Ina		
Volume of serial or monograph	137	Author of portion(s)	Sassenhagen, Jona; Schlesewsky, Matthias; Bornkessel-Schlesewsky, Ina
Page or page range of portion	29-39	Issue, if republishing an article from a serial	N/A
		Publication date of portion	2014-09-30