

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Event-related Potential (ERP) studies of the role of working memory, selective attention, and attentional efficiency in language acquisition and comprehension.

Permalink

<https://escholarship.org/uc/item/2n69z73n>

Author

Barkley, Christopher

Publication Date

2015

Supplemental Material

<https://escholarship.org/uc/item/2n69z73n#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Event-related potential (ERP) studies of the role of working memory, selective attention, and attentional efficiency in language acquisition and comprehension

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy

in

Linguistics and Cognitive Science

by

Christopher Michael Barkley

Committee in Charge:

Professor Robert Kluender, Chair
Professor Farrell Ackerman
Professor Terry Jernigan
Professor Marta Kutas
Professor Rachel Mayberry

2015

The dissertation of Christopher Michael Barkley is acceptable in quality and form for publication on microfilm and electronically.

Chair

University of California, San Diego

2015

TABLE OF CONTENTS

Signature Page.....	iii
Table of Contents.....	iv
List of Figures.....	v
List of Tables.....	vii
List of Supplementary Materials.....	viii
Acknowledgements.....	ix
Vita.....	x
Abstract of the Dissertation.....	xi
INTRODUCTION.....	1
CHAPTER 1: Referential Processing in the Human Brain: An Event-related Potential (ERP) Study.....	17
CHAPTER 2: The Role of Selective Attention in Implicit Language Learning: An Event-related Potential (ERP) Study.....	66
CHAPTER 3: Elicitation of Early Negativity in Sentence Processing Contexts Depends on Attentional Efficiency: An Event-related Potential (ERP) Study.....	130
CONCLUSION.....	178

LIST OF FIGURES

FIGURE 1.1: (A) Grand average ERP waveforms for object relative and coordinate clause controls. (B) Grand average ERP waveform for object relative and coordinate clause control conditions at electrode LLPf. (C) Topographic scalp iso-voltage map of the mean difference between conditions between 300-500 msec.....	30
FIGURE 1.2: (A) Grand average ERP waveforms for Pronoun, Co-referent and Pronoun, No Co-referent conditions. (B) Grand average ERP waveforms for both Pronoun conditions at electrode LMFr. (C) Topographic scalp iso-voltage map of the mean difference between conditions 300-500 msec. (D) Topographic scalp iso-voltage maps between 500 – 800msec.....	33
FIGURE 1.3: (A) Grand average ERP waveforms for the Name, Co-referent and Name, No Co-referent conditions (B) Grand average ERP waveforms for both Name conditions at channel LMFr.....	37
FIGURE 2.1: Two alternative-forced-choice accuracy over time during training session (all subjects).....	95
FIGURE 2.2: Two alternative-forced-choice reaction over time during training session (all subjects).....	96
FIGURE 2.3: Grammaticality judgment accuracy during testing session.....	97
FIGURE 2.4: (A) Grand average ERP waveforms for the grammatical controls, and agreement, transitivity, and auxiliary selection violations, for a one second epoch with 100 msec. baseline. (B) Grand average ERP waveforms for all conditions at electrode MiPf.....	100
FIGURE 2.5: Figure 2.5: (A) Topographic scalp iso-voltage map of the mean difference between (i) grammatical controls and agreement violations, (ii) between grammatical controls and transitivity violations, and (iii) between grammatical controls and auxiliary selection violations between 150-250 and 300-500 msec.....	101
FIGURE 2.6: Correlations between training accuracy and testing accuracy on grammatical sentences.....	103
FIGURE 2.7: Correlations between training accuracy and testing accuracy on agreement violations.....	103
FIGURE 2.8: Correlations between training accuracy and testing accuracy on auxiliary selection violations.....	104
FIGURE 2.9: Correlations between training accuracy and testing accuracy on transitivity violations.....	104
FIGURE 3.1: Schematic of the Attention Network Task (ANT).....	144
FIGURE 3.2: Mean RT for ANT performance.....	152
FIGURE 3.3: Plot of mean RT for ANT performance.....	153

FIGURE 3.4: All participants. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode RMFr between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.....	155
FIGURE 3.5: All participants. (A) Grand average ERP waveforms for Object Relatives and Controls (1 sec. epoch, 100 msec. baseline) between 300-500 msec.. (B) Grand average ERP waveform at electrode LLFr between 300-500 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	156
FIGURE 3.6: All participants. (A) Grand average ERP waveforms for Semantic Anomalies and Controls (1 sec. epoch, 100 msec. baseline) between 300-500 msec.. (B) Grand average ERP waveform at electrode LLFr between 300-500 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.....	157
FIGURE 3.7: Figure 3.6: All participants. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode MiPa between 500-800 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	158
FIGURE 3.8: Low efficiency alerting group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	160
FIGURE 3.9: High efficiency alerting group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.....	161
FIGURE 3.10: Low efficiency orienting group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	162
FIGURE 3.11: High efficiency orienting group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	163
FIGURE 3.12: Low efficiency executive group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	164
FIGURE 3.13: High efficiency executive group. (A) Grand average ERP waveforms for WCVs and Controls (1 sec. epoch, 100 msec. baseline) between 50-150 msec.. (B) Grand average ERP waveform at electrode LMPf between 50-150 msec, and (C) and topographic iso-voltage map of the difference between conditions in this time window.	165

LIST OF TABLES

TABLE 1.1: Experimental Sentences: Pronouns and proper names with and without preceding co-referents.....	28
TABLE 1.2: Experimental sentences: Object relative clauses and coordinate clause controls.....	28
TABLE 1.3: Filler sentences.....	56
TABLE 2.1: Example sentences of the artificial language used during training and testing.....	84
TABLE 3.1: Experimental sentences: Sentence-processing experiment.....	143

LIST OF SUPPLEMENTARY MATERIALS

FIGURE 1.4: Grand average ERP waveforms for object relative and coordinate clause controls at electrode LLPf. (A) *Multi-word analysis* (B) *Single-word analysis*. This analysis indicates that each sentence position does not make its own contribution to the effect seen in (A)

FIGURE 1.5: Grand average ERP waveforms for the Pronoun, Co-referent and Pronoun, No Co-referent at electrode LLFr. (A) *Multi-word analysis*. (B) *Single-word analysis*. This analysis indicates that each sentence position makes its own contribution to the effect seen in (A)

TABLE 3.3: Results of omnibus Analysis of Variance conducted on participant group as a whole as well as groups resulting from median splits on ANT performance. (H = High efficiency groups, L = Low efficiency groups)

TABLE 3.4: Results of distributional Analysis of Variance conducted on participant group as a whole as well as groups resulting from median splits on ANT performance

ACKNOWLEDGEMENTS

I would like to thank Dr. Robert Kluender for his seemingly unending support and guidance during the preparation of this dissertation.

I would also like to express my deep gratitude to Dr. Marta Kutas, and the members of the Language and Brain and Cognitive Electrophysiology labs for their advice and support throughout my years here.

Chapter 1, in full, is a manuscript submitted to *Brain Research*, and is currently under review. Barkley, C., Kluender, R., & Kutas, M (*submitted*). The dissertation author was the primary investigator and preparer of this manuscript.

Chapter 2, in full, is currently being prepared for submission for publication. Barkley, C., Kluender, R., & Kutas, M (*in preparation*). The dissertation author was the primary investigator and preparer of this manuscript.

Chapter 3, in full, is currently being prepared for submission for publication. Barkley, C., Kluender, R., & Kutas, M (*in preparation*). The dissertation author was the primary investigator and preparer of this manuscript.

VITA

Bachelor of Arts, Linguistics, University of Florida, 2005

Master of Arts, Linguistics, University of Florida, 2007

Doctor of Philosophy, Linguistics and Cognitive Science, 2015

PUBLICATIONS

Kaan, E., Wayland, R., Bao, M. and Barkley, C. (2007). Effects of native language and training on lexical tone perception: an ERP study. *Brain Research* 1148, 113-122.

Kaan, E., Dallas, A.C., and Barkley, C. (2007). "Processing Bare Quantifiers in Discourse." *Brain Research* 1146, 199-209.

Kaan, E., Barkley, C., Bao, M., and Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: an event-related potentials training study. *BMC Neuroscience* 9.

Wayland, R., Kaan, E., Bao, M., and Barkley, C. (2008). Effects of native language and training on speaker normalization on lexical tone perception: a behavioral and ERP study. *Kansas Working Papers in Linguistics* 30, 359-372.

Kluender, R., Kutas, M., and Barkley, C. (2012). On Beyond Violations: Progress in Linguistic ERP Research. In: Festschrift for Marica De Vincenzi (*submitted*).

Barkley, C., Kluender, R., and Kutas, M. (2015). The Cognitive Processes Involved in Referential Dependency Formation (*forthcoming*).

Barkley, C., Kluender, R., and Kutas, M. (2015). Elicitation of Early Negativity (EN) in Sentence Processing Contexts Depends on Attentional Efficiency (*in preparation*).

Barkley, C., Kluender, R., and Kutas, M. (2015). The Role of Selective Attention in Artificial Grammar Learning (*in preparation*).

FIELDS OF STUDY

Linguistics
Cognitive Science
Neuroscience

ABSTRACT OF THE DISSERTATION

Event-related potential (ERP) studies of the role of working memory, selective attention, and attentional efficiency in language acquisition and comprehension

by

Christopher Michael Barkley

Doctor of Philosophy in Linguistics and Cognitive Science

University of California, San Diego, 2015

Professor Robert Kluender, Chair

Language researchers have long been interested in the extent to which performance factors play a central role in language use, and how these factors interact with linguistic representations of the competence grammar. Here I address this issue in three ERP experiments, investigating the role that domain-general neuro-cognitive systems play in language acquisition and comprehension.

In Experiment 1, I test the hypothesis that the formation of both long-distance syntactic dependencies and referential dependencies is underpinned by the same cognitive operations. I show that second elements in each dependency type elicit

brain responses associated with the domain-general working memory system. Though the constraints on the formation of these two dependency types have traditionally been treated as distinct in the theoretical literature, I show that the same basic working memory processes are operative in each case. Nonetheless, subtle differences in the brain responses elicited are consistent with construction-specific properties of these dependencies as well.

In Experiment 2, I employ a traditional training-testing paradigm to show that selective attention plays a subtle, nuanced role in implicit language learning. After an hour of training, the brain showed domain-general responses consistent with implicit learning of grammatical systems unattended during training, though behavioral performance on judging the grammaticality of violations of these systems remained below chance, suggesting that the representations generated under brief implicit learning conditions are not sufficiently robust to influence behavior. I discuss these results in the context of studies of both implicit learning of non-linguistic information and those investigating second-language learning of grammatical and semantic information

Lastly, in Experiment 3, I utilize a sentence-processing experiment and a task that measures attentional efficiency and show that the elicitation of early negativity (or eLAN) in sentence-processing contexts depends to a large degree on attentional efficiency, and therefore that these brain responses are not underlyingly linguistic, but should rather be interpreted as attentional modulations of domain-general N100 responses.

In general, I contend that the notion of “linguistic ERP components” may be misguided, and that what initially appear as language-specific responses should rather be interpreted as indices of domain-general cognitive systems operating over structured, statistically rich linguistic input.

INTRODUCTION

Early linguistic theorizing, based on the pioneering work of Noam Chomsky in the mid 20th century, posited a fully specified, highly articulated, and innately specified Universal Grammar to account for the acquisition of linguistic competence, i.e. the knowledge of a language possessed by a native speaker, and the subsequent use of this knowledge. This detailed linguistic competence, argued to be largely invariant across speakers, was claimed to be independent of *performance* factors (such as memory and attention), in other words the cognitive factors that constrain the efficiency of language use. These factors, more recently referred to as “third factors,” (see below), *do* vary considerably across individuals, and can account for the large amount of variability in language behaviors. However, these performance factors were largely relegated to the periphery in this theoretical framework. In this dissertation I aim to show that such factors are more central to the architecture of the language system and language behaviors than was originally assumed¹ and fact, I will argue that the cognitive systems underlying these performance factors actually form hard-wired interfaces with the language system. Proponents of the theory of Universal Grammar argue that certain (core) aspects of grammatical knowledge are innate and therefore constitute a part of our biological endowment. Under this original set of nativist assumptions, this genetically encoded knowledge is central to the process via which language users move from a start state, through the process of language acquisition, to an end state in which they possess full knowledge of their native language(s).

¹ It should be noted however that the importance of these factors has recently given a more central role in the literature, as discussed below.

In this view, the cognitive operations performed over linguistic representations are argued to be *domain-specific*. In the context of this terminology, “domain” simply refers to a type of input (language, visual stimuli, tactile stimulation, etc.), and “specific” refers to any process whose operations are restricted to a single input domain (in this case, language). Early arguments in support of this position were that the language input present in the environment is too sparse and perceptually degraded for language to be successfully acquired (the *poverty of the stimulus* argument), that there is a lack of negative evidence to guide language learning (parents rarely correct their children’s errors), and that children acquire language rapidly and effortlessly with minimal instruction. In sum, the claim is that language is simply too complex a cognitive system to be learned and therefore there must be *something* encoded in the genome, that functions to guide the process of language acquisition.

While in recent years strong version of this position has begun to fall out of favor, there are still many researchers who adhere to (at least some form of) the theory of Universal Grammar (e.g. Boeckx & Grohmann, 2007; Boeckx & Longo, 2011; Boeckx, 2013). However, during the last decade of the 20th century, a substantial amount of research, discussed in detail below, began to emerge that questioned the extent to which language was in fact un-learnable, as had previously been asserted. To be clear, if it in fact could be shown that language is learnable, then there would be no (or at least less of a) need for the theory of Universal Grammar to account for the acquisition of linguistic competence. This research program demonstrated the existence of powerful and flexible mechanisms that could be utilized during the learning of any complex and statistically rich system (including language) containing learnable

regularities, and therefore could account for language acquisition without recourse to the existence of highly specified linguistic knowledge. Put simply, this research suggests that language *is* learnable. These accounts are consistent with the notion that the acquisition of linguistic competence and the application of this knowledge during language use can be accounted for with *domain-general* mechanisms (“general” in this sense referring to mechanisms whose operations are not performed over a single input domain). These mechanisms are not specific to language but are applied in a general fashion across cognitive (and sensorimotor) domains as we learn from the world by taking advantage of its complex structure.

While all of these learning mechanisms and process of (linguistic) data analysis and the constraints on their application have no doubt been incompletely described, evidence for a number of them has been comprehensively demonstrated and replicated across studies. Examples of these abilities include statistical/distributional learning at multiple linguistic levels (Saffran, 1996; Saffran, 2001; Saffran, 2002; Saffran & Wilson, 2003; Thompson & Newport, 2007; Wonnacott et al., 2008), categorical perception (Eimas et al., 1971; Jusczyk et al., 1999), and application of the mutual exclusivity constraint (Markman & Wachtel, 1988). Additional support for the domain-general nature of these learning mechanisms comes from observations that they can be engaged during the learning of *non-linguistic* information: categorical perception has been demonstrated with tonal stimuli (Creel et al., 2004) as well as during the processing of color information and facial expressions (Roberson et al., 1999). Furthermore statistical learning of visual shape information has been observed in many studies (e.g. Fiser & Aslin, 2001; Fiser et al., 2002). In addition, the existence of these

learning mechanisms has been documented in other species, providing further evidence that they are not specific to language. It has been shown that cotton-top tamarins (Hauser et al., 2001) can use statistical cues to segment “words” from a continuous speech stream, that both chinchillas (Kuhl & Miller, 1978) and cotton-top tamarins have the ability to categorically perceive continuous linguistic stimuli (Ramus et al., 2000), and that canines have the ability to apply the mutual exclusivity constraint to learn the referents of words (Kaminski et al., 2004) often after single exposure to the word (“fast mapping”). Perhaps even more strikingly, recently researchers have observed rudimentary combinatorial abilities, long thought to be unique to human language, in Campbell’s monkeys (Ottarrra et al., 2009).

In sum, these observations demonstrate that (i) at least some aspects of language are learnable via the application of powerful domain-general learning mechanisms, (ii) that these learning mechanisms also operate over non-linguistic input, and (iii) that the rudiments of these abilities can be found in non-human species. As stated above, these observations have led many researchers to argue that there is no need to posit the existence of species-specific genetically encoded linguistic knowledge to explain the emergence of linguistic competence, if this competence can in fact be derived from the actions of general learning mechanisms operating over linguistic input. In other words, it appears that language is a new machine built out of old “parts.”

In more recent work, proponents of the theory of Universal Grammar have begun to refer to these types of mechanisms as “third factors” (see for example Boeckx, 2008; Chomsky, 2005; Trotzke et al., 2013). These theorists have argued that these third factors are “computational systems not specific to language” (Trotzke et al., 2013:3),

and “principles of data analysis that might be used in language acquisition and in other domains [...] including principles of developmental constraints” (Chomsky, 2005:6). Following from these claims, it has been argued that the role of these third factors has “led to a much less specified view of the genetic endowment of language (UG)” (Trotzke et al., 2013:3), and therefore “an architecture of grammar that is more plausible biologically than a fully specified, highly specific UG” (Boeckx, 2008: 14). Nonetheless, even with a developing account of the interaction between UG and these third factors, (language specific) principles such as recursive (internal and external) Merge, binary branching structures, the valued-undervalued distinction, and the distinction between heads and complements are still argued to be part of UG and therefore innately specified. Though this position no doubt represents an improvement over initial descriptions of the contents of an innately specified Universal Grammar, it remains the case that supporters of this position still contend that language-specific information is innately specified.

Below I discuss Event-Related Potential (ERP) studies of language processing, the results of which are relevant to the debate as to the domain-specific or domain-general mechanisms underlying language use. The results of the experiments described in this dissertation bear directly on this issue, providing further support for domain-general accounts of language acquisition and on-line processing of linguistic input.

Ever since the seminal work of Marta Kutas and Steven Hillyard in the early 1980s (e.g. Kutas & Hillyard 1980, 1983), there has been an explosion of ERP research investigating online language comprehension (see Kutas et al., 2006 for a comprehensive review of language-related ERP research). During an ERP experiment,

participants are presented with linguistic stimuli while their continuous electroencephalogram (EEG) is recorded. The EEG reflects the synchronous electrical activity of thousands of spatially aligned pyramidal neurons and the resulting post-synaptic field potentials generated by this activity. After the completion of the experiment, the brain's response to multiple instances of the same experimental condition is averaged, thus eliminating noise unrelated to linguistic processing, and yielding a series of positive and negative voltage deflections. The resulting waveforms are known as the ERP. These voltage deflections, time-locked to the stimulus of interest, are referred to as components, which differ in terms of their polarity, peak latency, and distribution at the scalp. ERPs are thus a multi-dimensional measure, and differences in the parameters of the brain responses elicited by the experimental conditions of interest are taken as an indication that the brain is sensitive to the linguistic information that is being experimentally manipulated. Because the activity of an infinite number of combinations of underlying source-generator configurations can combine to yield the same pattern of electrical activity at the scalp, it is impossible to make a backwards (spatial) inference from voltage deflections seen at the scalp to the brain areas responsible for generating these signals (the "inverse problem"), without conducting additional analyses. However, the high degree of temporal resolution, at the level of milliseconds, makes the ERP technique ideal for studying real-time language processing.

During the early years of linguistic ERP research, researchers focused primarily on semantic (e.g. Kutas & Hillyard, 1981) and syntactic (e.g. Neville et al., 1991; Friderici et al., 1993; Osterhout & Holcomb, 1992) processes and their attendant

electrophysiological indices, the N400 and *e*LAN/LAN/P600, respectively. Focusing on issues of modularity (Fodor, 1983), many studies also investigated the extent to which these processes could be dissociated (e.g. Münte et al., 1993). However, recent research has made much progress in terms of mapping the antecedent conditions under which language-related ERP components are elicited,² as well as going beyond the violation paradigms that initially dominated the field to study other factors that influence language comprehension and observed brain responses. To name but a few of these research areas, much work has been done on the processing of referential relationships within and across sentences (e.g. Anderson & Holcomb, 2005; Heine, 2006; Nieuwland & Van Berkum, 2006; Nieuwland et al., 2007; Osterhout & Mobley, 2007; Van Berkum et al., 1999, 2003, 2004, 2007; Nieuwland, 2014), the role of top-down discourse-level influences on the processing of individual words (Van Berkum, J., Hagoort, P., & Brown, C. 1999; Nieuwland, M., & Van Berkum, J. 2006 and Filik, R., & Leuthold, H. 2008), the influence of extra-linguistic factors such as mood (Egidi & Nussbaum, 2012) and speaker gender (Van Berkum et al., 2008; 2009; Regel et al., 2010), and the role of the right hemisphere in the processing of non-literal language (Coulson & Van Petten, 2007; Wlotko & Federmeier, 2007; Federmeier & Wlotko, 2008). These findings further underscore the utility of the ERP technique as a tool to study language comprehension, as they have contributed greatly to our understanding of the linguistic and extra-linguistic factors that influence language processing and the temporal structure of the language-processing stream.

² For example, see Kutas & Federmeier (2011) for a comprehensive discussion of the variables that modulate the N400 response and the underlying cognitive processes that the component is assumed to index.

Based on the preceding discussion, it is clearly the case that linguistic ERP research has made a great deal of progress in mapping the antecedent conditions under which linguistic ERP responses are elicited, as well as the linguistic factors that specific components are sensitive too. However, an issue that has persisted in the field is the extent to which these brain responses to linguistic stimuli index domain-specific or domain-general processes. Recall that simply because linguistic stimuli reliably elicit these responses, it is not necessarily the case that they constitute responses to language *per se*, but rather could reflect more general cognitive responses triggered by any form of structured input. In fact, for every linguistic component thus far documented, one could argue that this appears to be the case.

Domain-general accounts have been proposed for N400 and LAN/P600 effects that have long been interpreted as indexing the semantic and syntactic aspects of language processing, respectively. The N400 is now often interpreted as reflecting the processing of linguistic *and* non-linguistic meaning (see Sitnikova et al., 2008 for a review), with N400 responses elicited, and modulated by, manipulations involving pictures (Ganis et al., 1996), the structure of visual narratives (Cohn et al., 2011; Cohn 2014), videos (Sitnikova et al., 2003), faces (Olivares et al., 1999), environmental sounds (Van Petten & Rheinfelder 1995), gestures (Wu & Coulson, 2005, 2007, 2010), and mathematical sequences (Niedeggen et al., 1999). There has been a long-standing argument about the domain-specificity of the P600 often observed in response to (morpho-)syntactic violations (e.g. Coulson et al., 1998; Osterhout, 1999), a response that is also seen when processing violations of structured musical (Besson & Macar, 1987, Patel et al., 1998) and geometric (Besson & Macar, 1987) sequences. Even

within the linguistic domain, recent research has show that P600 effects can be elicited by violations that are not purely syntactic in nature, but rather appear to arise when difficulties assigning thematic roles are encountered. These effects have been argued to reflect conflicts between semantic and combinatorial processing streams (Kuperberg, 2007), thematic and pragmatic processing streams (Bornkessel & Schlesewsky, 2006), or domain-general conflict occurring outside of the language system (Kolk & Chwilla, 2007) Likewise, left anterior negativities (LAN) observed during the processing of morpho-syntactic violations (e.g. Kutas & Hillyard, 1983; Neville et al., 1991), long-distance dependencies, and non-canonical word order have been argued to reflect the burdens imposed by these configurations on working memory (or some other pool of cognitive resources, see Martin-Loeches et al., 2005), rather than indexing linguistic operations *per se* (see, among others, Kluender & Kutas, 1993; King & Kutas, 1995; Munte et al., 2008 ; Rösler et al, 1998; Matzke et al., 2002; Ueno & Kluender, 2003; Hagiwara et al., 2007; Ueno & Garnsey 2007; and Kwon et al., 2013). In further support of this position, the amplitude of LAN effects has been shown to co-vary with an individual's working memory score as assessed by the reading-span task (Daneman & Carpenter, 1980; see Vos et al., 2001; Munte et al., 1998; Gunter et al., 2003; Fiebach, et al., 2004; Nieuwland & Van Berkum, 2006 for examples). Lastly, the component originally known as the early left anterior negativity (eLAN), argued to index phrase-structure building processes occurring during the initial stage of a modular and serial parser, has recently been argued to index processes that are not syntactic or even linguistic in nature, but rather sensory responses to low-level physical features of linguistic input (Dikker, 2009; Dikker et al., 2010, 2011).

In the context of these issues, the contributions of this dissertation are two-fold. I provide additional support for domain-general characterizations of language acquisition and processing, and the results of the experiments described herein are consistent with accounts of language use that emphasize the central role of performance factors. Secondly, I add to an accumulating series of observations that suggest that the notion of a “linguistic ERP component” may be misguided, and these effects may rather index domain-general cognitive processes in response to linguistic input. In the following chapter I discuss an experiment investigating the formation of long-distance antecedent-pronoun relationships. The results show that brain responses (LAN) are elicited in response to the second elements in these dependencies that are similar to those observed in response to the second element in long-distance syntactic dependencies. Referential and syntactic dependencies have long been kept distinct in the theoretical linguistics literature, with a debate centering on which dependency types are handled exclusively by the syntax, and which are handled via a combination of syntactic constraints and discourse principles (which may or may not be innately specified) (Chomsky, 1980, 1981; Reinhart, 1983; Pesetsky, 1987; Cinque, 1990; Chung, 1994). The fact that the formation of each of these dependency types elicits similar brain responses, arguably indexing working memory-based operations, further articulates and emphasizes the role of domain-general cognitive systems in online language comprehension. Chapter 2 contains a discussion of the results of an experiment investigating the role of selective attention in implicit language learning. This study shows that only grammatical systems that are attended during learning produce representations sufficiently robust to influence behavior, though there was

some evidence of learning in the form of brain responses to violations of the unattended systems. As such, these results suggest that low-level attentional processes play a nuanced but crucial role during language acquisition. Lastly, in Chapter 3 I report an experiment showing that the elicitation of *e*LAN, originally argued to index phrase-structure building operations, depends on attentional efficiency, as was only participants with low-efficiency attentional networks showed evidence of this response. I conclude by interpreting the sum of the results of all three experiments in the context of the issues outlined above.

Works cited

- Besson, M., & Macar, F. (1987). An Event-Related Potential Analysis of Incongruity in Music and Other Non-Linguistic Contexts. *Psychophysiology*, 24, 14-25.
- Boeckx, C., & Grohmann, (2007). The Bio-linguistic manifesto. *Biolinguistics*, 1, 1-8.
- Boeckx, C. & Longa, V. (2011). Lenneberg's Views on Language Development and their Relevance for Modern Biolinguistics. *Biolinguistics*, 5, 254-273.
- Boeckx (2013). Biolinguistics: Forays into Human Cognitive Biology. *Journal of Anthropological Sciences*, 91, 63-89.
- Bornkessel-Schlesewsky, I., Schlewsky, M., (2008). An Alternative Perspective on Semantic P600 Effects in Language Comprehension. *Brain Research Review*. 59, 55–73.
- Chomsky, N. (1980). On Binding. *Linguistic Inquiry*, 11(1), 1-46.
- Chomsky, N. Lectures on Binding and Government. Foris Publications, Dordrecht. 1981.
- Chung, S. (1994). Wh-agreement and Referentiality in Chamorro. *Linguistic Inquiry*, 25, 1-44.
- Cinque, G. (1990). Types of A'-dependencies. Cambridge, MA: MIT Press
- Cooke, A. Zurif, E., DeVita, C., Alsop, D., Koenig, P, Detre, J., Gee, J, Pinango, N., Balogh, J. & Grossman, M. (2001). Neural Basis for Sentence Comprehension: Grammatical and Short-term Memory Components. *Human Brain Mapping*, 15, 80-94.
- Cohn, Neil, Ray Jackendoff, Phillip Holcomb, and Gina Kuperberg. (2011). The grammar of visual narratives: Neural evidence for constituent structure in visual narrative comprehension. *Neuropsychologia*. 64, 63-70.

- Cohn, Neil. (2014). You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science*, **38**(7), 1317-1359.
- Coulson, S. King, J., & Kutas, M. (1999). ERPs and Domain Specificity: Beating a Straw Horse. *Language and Cognitive Processes*, **13**, 653-672.
- Coulson, S., & Van Petten, C. (2007). A Special Role for the Right Hemisphere in Metaphor Comprehension?: ERP evidence from hemifield presentation. *Brain Research*, 1146, 128-145.
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant Melodies: Statistical Learning of Non-adjacent Dependencies in Tone Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 1119-1130.
- Daneman, M., & Carpenter, P. (1980). Individual Differences in Working memory and Reading. *Journal of Verbal Learning and Verbal behavior*. **19**, 450-466.
- Dikker, S., Rabagliatti, H., Pykkänen, L. (2009). Sensitivity to Syntax in Visual Cortex. *Cognition* **110** (3), 293-321.
- Dikker, S., Rabagliatti, H., Farmer, T., & Pykkänen. (2010). Early Occipital Sensitivity to Syntactic category is Based on Form Typicality. *Psychological Science*, **21**(5), 629-634.
- Dikker, S., & Pykkänen, L. (2011). Before the N400: Effects of Lexico-Semantic Violations in Visual Cortex. *Brain and Language*, **118**(1-2), 23-28.
- Egidi, G., & Nussbaum, H. (2012). Emotional Language Processing: How Mood Affects Integration Processes during Discourse Comprehension. *Brain & Language*, **122**, 199-210.
- Eimas, Siqueland, Jusczyk, & Vigorito (1971). Speech Perception in Infants. *Science*, **171**, 303-306.
- Federmeier, K., Wlotko, E., & Meyer, E. (2008). What's 'Right' in Language Comprehension: Event-Related Potentials Reveal Right Hemisphere Language Capabilities. *Language and Linguistics Compass*, **2**, 1-17.
- Fiebach, C., Schesewsky, M., & Friederici, A. (2004). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. *Journal of Memory and Language*, **47**, 250-272.
- Filik, R., & Leuthold, H. (2008). Processing Local Semantic Anomalies in Fictional Contexts: Evidence from the N400. *Psychophysiology*, **45**, 554-558.
- Fiser, J., & Aslin, R. (2001). Unsupervised Statistical Learning of Higher-order Spatial Structures from Visual Scenes. *Psychological Science* **12**, 499-504.
- Fiser, J., & Aslin, R. (2002). Statistical Learning of Higher-order Temporal Structure from Visual Shape Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **28**(3), 458-467.
- Fodor, JA. The Modularity of Mind. Cambridge, Massachusetts: Bradford Publishing, 1981. Print.
- Friederici, AD., 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.

- Ganis, G., Kutas, M., & Sereno, M. (1996). The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Neuroscience*, **8**, 89-106.
- Gunter, T., Wagner, S., & Friederici, A. (2003). Working Memory and Lexical Ambiguity Resolution as Revealed by ERPs: A Difficult Case for Activation Theories. *Journal of Cognitive Neuroscience*, **15**, 643-657.
- Hagiwara, H., Soshi, T., Isihara, M., & Imanaka, K. (2007). A Topographical Study on the Event-Related Potential Correlates of Scrambled Word Order in Japanese Complex Sentences. *Journal of Cognitive Neuroscience*, **19**(2), 175-193.
- Hauser, M., Newport, E., & Aslin, R. (2001). Segmentation of the Speech Stream in a Non-human Primate: Statistical Learning in Cotton-top Tamarins. *Cognition*, **78**, B63-B64.
- Heine, A., Tamm, S., Hoffman, M., Hutzler, F., and Jacobs, A. (2006). Does the Frequency of the Antecedent Noun Affect the Resolution of Pronominal Anaphors? An ERP Study. *Neuroscience Letters*, **400**, 7-12.
- Jusczyk, P., Houston, D.M., & Newsome, M. (1999). The Beginnings of Word segmentation in English-Learning Infants. *Cognitive Psychology*, **39**, 159-207.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word Learning in a Domestic Dog: Evidence for “Fast Mapping.” *Science*, **304**, 1682-1672.
- King, J., & Kutas, M. (1995). Who Did What and When? Using Word- and Clause-Level ERPS to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, **7**(3), 376-395.
- Kluender, R., & Kutas, M. (1993a). Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. *Journal of Cognitive Neuroscience*, **5**(2), 196-214.
- Kolk, H.H.J., Chwilla, D.J., 2007. Late Positivities in Unusual Situations. *Brain and Language*, **100**, 257-261.
- Kuhl, P., & Miller, J. (1978). Speech Perception by the Chinchilla: Identification Functions for Synthetic VOT stimuli. *The Journal of the Acoustic Society of America*, **63**, 905.
- Kuperberg, G. (2007). Neural Mechanisms of Language Comprehension: Challenges to Syntax. *Brain Research*, **114**, 23-49.
- Kutas, M., & Hillyard, S. (1981). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, **207**, 203-205.
- Kutas, M., & Hillyard, S. (1983). Event-related Potentials to Grammatical Errors and Semantic Anomalies. *Memory and Cognition*, **11**, 539-550.
- Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the Eventrelated Brain Potential (ERP). *Annual Review of Psychology*, **62**, 621-647.
- Kwon, N., Kluender, R., Kutas, M., & Polinsky, M. (2013). Subject/object Processing Asymmetries in Korean Relative Clauses: Evidence from ERP data. *Language*, **89**(3), 537-585.
- Markman, E., Wachtel, G. (1988). Children’s Use of Mutual Exclusivity to Constrain the Meaning of Words. *Cognitive Psychology*, **20**, 121-157.

- Martin-Loeches, M., Munoz, F., & Casado, P. (2005). Are the Anterior Negativities to Grammatical Violations Indexing Working Memory? *Psychophysiology*, **42**(5), 508-519.
- Matzke, M., Mai, H., Nager, W., Russeler, J., & Münte, T. (2002). The Costs of Freedom: An ERP Study of Non-canonical Sentences. *Clinical Neurophysiology*, **113**, 844-852.
- Munte, T., Heinze, H., & Mangun, G. (1993). Dissociation of Brain Activity Related to Syntactic and Semantic Aspects of Language. *Journal of Cognitive Neuroscience*, **5**, 335-354.
- Münter, T., Schlitz, S., & Kutas, M. (1998). When Temporal Terms Belie Conceptual Order. *Nature*, **395**, 71-74.
- Neville, H.J., Nicole, J., Barss, A., Forster, K., & Garret, M. (1991). Syntactically-based Sentence Processing Classes: Evidence From Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, **3**, 155-170.
- Niedeggen, Rösler, M., Jost, K. (1999). Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. *Psychophysiology*, **36**, 307-324.
- Nieuwland, M., & van Berkum, J. (2006). Individual Differences and Contextual Bias in Pronoun Resolution. *Brain Research*, **1118**, 155-167.
- Nieuwland, M., & van Berkum, J. (2006). When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. *Journal of Cognitive Neuroscience*, **18**, 1098-1111.
- Nieuwland, M., Petersson, K.M., & Van Berkum, J. (2007). On Sense and Reference: Examining the Functional Neuroanatomy of Referential Processing. *NeuroImage*, **37**, 993-1004.
- Nieuwland, M., Otten, M., & Van Berkum, J. (2007). Who Are You Talking About? Tracking Discourse Level Referential Processing with Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, **19**(2), 228-236.
- Nieuwland (2014). "Who's he?" Event-related Brain Potentials and Unbound Pronouns. *Journal of Memory and Language*, **(76)**, 1-28.
- Olivares, E., Iglesias, J., & Rodríguez-Holguín, S. (1999). Long-Latency ERPs and Recognition of Facial Identity. *Journal of Cognitive Neuroscience*, **13**, 136-151.
- Osterhout, L., & Holcomb, P. (1992). Event-related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language*, **31**(6), 785-806.
- Osterhout, L., & Mobley, L. (1995). Event-related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*, **34**(6), 739-773.
- Osterhout, L. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, **14**, 1-14.
- Ouataara, K., Lemasson, A., & Klaus, Z. (2009). Campbell's Monkeys Concatenate Vocalizations into Context-specific Call Sequences. *Proceedings of the Natural Academy of Science*, **106**, 22026-22031.
- Patel, A., Gibson, T., Ratner, J. & Besson, M. (1998). Processing Syntactic Relations in Language and Music: An Event-Related Potential Study. *Journal of Cognitive Neuroscience*, **10**, 717-733.

Pesetsky, D. (1990). Wh-in-situ: Movement and Selective Binding. In Reuland and Martin (eds.), *The representation of (in-)definiteness*. Cambridge, MA: MIT Press.

Ramus, F., Hauser, M., Miller, C., Morris, J., & Mehle, J. (2000). Language Discrimination by Human Newborns and by Cotton-top Tamarin Monkeys. *Science*, **288**, 349-351.

Regel, S., Coulson, S., & Gunter, T. (2010). The Communicative Style of a speaker can Affect Language Comprehension? ERP evidence from the Comprehension of Irony. *Brain Research*, **1311**, 121-135.

Reinhart, T. (1983). Coreference and Bound Anaphora: A Restatement of the Anaphora Questions. *Linguistics and Philosophy*, **6**, 47-88.

Roberson, D., Davidoff, J., & Brainsby, N. (1999). Similarity and Categorisation: Neuropsychological Evidence for a Dissociation in Explicit Categorization Tasks. *Cognition*, **71** 1-42.

Rösler, F., Pechmann, T., Streb, J., Roder, B., & Henninghausen (1998). Parsing of Sentences in a Language with Varying Word Order: Word-by-word Variations of Processing Demands are Revealed by Event-Related Potentials. *Journal of Memory and Language*, **38(2)**, 150-176.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical Learning by 8-month-old Infants. *Science*, **274**, 1926-1928.

Saffran, J. (2001). The Use of Predictive Dependencies in Language Learning. *Journal of Memory and Language*, **44**, 493-515.

Saffran, J. (2002). Constraints on Statistical Language Learning. *Journal of Memory and Language*, **47**, 172-196.

Saffran, J. and Wilson, D. (2003). From Syllables to Syntax: Multilevel Statistical Learning by 12-month-old infants. *Infancy*, **4**, 273-284.

Sitnikova, T., Kuperberg, G., & Holcomb, P. (2003). Semantic Integration in Videos of Real-world events: An Electrophysiological Investigation. *Psychophysiology*, **40**, 160-164.

Sitnikova, T., Holcomb, P., & Kiyonaga, K. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, **20**, 237-257.

Thompson, S., & Newport, E. (2007). Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning and Development*, **3**, 1-42.

Ueno, M., & Kluender, R. (2003). Event-related Brain Indices of Japanese Scrambling. *Brain and Language*, **86**, 243-271.

Ueno, M. & Garnsey, S. (2007) An ERP Study of the Processing of Subject and Object Relative Clauses in Japanese. *Language and Cognitive Processes*, **23(5)**, 646-688.

Van Berkum, J., Brown, C., & Hagoort, P. (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials. *Journal of Memory and Language*, **41**, 147-182.

Van Berkum, J., Brown, C., & Hagoort, P. (1999). Semantic Integration in Sentences and Discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, **11**, 657-671.

Van Berkum, J., Brown, C., Hagoort, P., & Zwisterlood, P. (2003). Event-Related Brain Potentials Reflect Discourse-Referential Ambiguity in Spoken Language Comprehension. *Psychophysiology*, **40**, 235-248.

Van Berkum, J.J.A. van (2004). Sentence Comprehension in a Wider Discourse: Can we Use ERPs to Keep Track of Things? In M. Carreiras, Jr. & C. Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York: Psychology Press.

Van Berkum, J., Koorneef, A., Otten, M., & Nieuwland, M. (2007). Establishing Reference in Language Comprehension: An Electrophysiological Perspective. *Brain Research*, **1146**, 158-171.

Van Berkum, J., van Den Brink, D., Tesink, C., Kos, M., & Hagoort, P. (2008). The Neural Integration of Speaker and Message. *Journal of Cognitive Neuroscience*, **20**, 580-591.

Van Berkum, J., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or Wrong?: The Brain's Fast Response to Morally Objectionable Statements. *Psychological Science*, **20**, 1092-1099.

Vos, S., Gunter, T., & Schriefers, H. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, **16**, 65-103.

Petten, C., & Rheinfelder, (1999). Conceptual Relationships between Spoken Words and Environmental Sounds: Event-related Brain Potential Measures. *Neuropsychologia*, **33**, 454-508.

Wlotko, E., & Federmeier, K. (2007). Finding the Right word: Hemispheric Asymmetries in the Use of Sentence Context Information. *Neuropsychologia*, **45**, 3001-3014.

Wonnacott, E., Newport, E., & Tanenhaus, M. (2008). Acquiring and Processing Verb Argument Structure: Distributional Learning in a Miniature Language. *Cognitive Psychology*, **56**, 165-209.

Wu, Y.C. & Coulson, S. (2005). Meaningful Gestures: Electrophysiological Indices of Iconic Gesture Comprehension. *Psychophysiology* **42**: 654-667.

Wu, Y.C. & Coulson, S. (2007). How Iconic Gestures Enhance Communication: An ERP Study. *Brain & Language* **101**:234-245.

Wu, Y., & Coulson, S. (2010). Gestures modulate speech processing early in utterances, *Neureport*, **21**, 521-526.

CHAPTER 1: Referential Processing in the Human Brain: An Event-related Potential (ERP) Study

Abstract:

A substantial body of ERP research investigating the processing of syntactic long-distance dependencies has shown that, across languages and construction types, the second element in such configurations typically elicits phasic left anterior negativity (LAN). We hypothesized that these effects are not specific to syntactic dependencies, but rather index a more general cognitive operation in which the second (dependent) element in sentence-level linguistic long-distance relationships triggers a process of association with the first element. We tested this hypothesis with straightforward referential dependencies, comparing pronouns with proper name antecedents to those without, and proper names with and without preceding co-referring pronouns. We predicted phasic LAN effects in response to the second referential element in both comparisons, but observed them only in response to pronouns with antecedents; no differences were observed between responses to proper names with and without preceding co-referring pronouns. We argue that LAN effects observed at the pronoun index the cognitive operations necessary for the association of a pronoun with its antecedent, on which it depends for its reference. Similar but non-identical responses were elicited by the main clause verb following the gap position in object relative clause constructions compared to coordinate clause controls in an orthogonal manipulation. LAN effects were thus elicited by the second dependent element in both construction types, suggesting that long-distance syntactic and referential dependencies pose similar processing challenges. These findings help to clarify the cognitive processes indexed by

anterior negative responses to associated dependent elements in a variety of language contexts.

1. Introduction

1.1 Long-distance dependencies

A pervasive property of human language that poses a great challenge to the comprehension system is that non-adjacent sentence elements often depend on each other for successful interpretation. This interpretation process may require the formation of interdependent relationships between such elements across arbitrarily long distances.¹ Modern linguistic theory has traditionally divided sentence-level long-distance dependencies into two main types, which we refer to here as “syntactic” and “referential.” These have long been the subject of debates as to exactly which dependency types are handled exclusively by the syntax, and which are handled via a combination of syntactic constraints and discourse principles (Chomsky, 1980, 1981; Reuland & Reinhart, 1993; Pesetsky, 1987; Cinque, 1990; Chung, 1994).

English relative clauses are a common and oft-studied example of a syntactic long-distance dependency. In these constructions, a sentence constituent like *The proposal* in (1) appears at the left edge of its clause, rather than in its underlying or “base” position, where it would ordinarily appear in a simple declarative clause, indicated by *the proposal* in (2) and an underscore in (1).

(1) *The proposal* [that they’re currently in the process of finalizing ____] seems solid.

(2) They’re currently in the process of finalizing *the proposal*.

¹ Here we restrict ourselves to a discussion of sentence-level non-local relationships, while fully recognizing the existence of word-internal non-local relationships, such as phonological processes of vowel harmony (Rose & Walker, 2011) and long-distance consonant assimilation (Rose, 2011).

In the psycholinguistic and neurolinguistic literatures, the displaced element (*the proposal* in (1)) is referred to as the “filler” and the underscore in (1) is referred to as a “gap”. Forming a relationship between filler and gap, a so-called “filler-gap dependency” (Fodor, 1978), is necessary for comprehension of relative clauses like (1) because the filler must be interpreted as the direct object of the verb *finalize* and the undergoer of the *finalizing* action.

A common example of a referential dependency is the relationship between a pronoun and its antecedent, as in (3). Referential dependencies do not involve the dislocation of any sentence constituent, but do require the formation of a relationship between two non-local elements (the pronoun *it* in and its antecedent *the proposal* in (3)) for comprehension to succeed.

(3) *The proposal* is currently being finalized and *it* seems solid.

In this paper, we are concerned with the brain responses elicited by the second of the two elements in such long-distance dependencies (the gap position in (1) and the pronoun *it* in (3)). Specifically, we aim to determine the extent to which the brain treats syntactic and referential dependencies similarly or differently, as indexed by the electrical brain responses elicited by the second elements in these two types of long-distance relationships.

We first review what is known about the brain’s response to the second element in syntactic dependencies, focusing on studies using event-related brain potentials (ERPs). The precise temporal resolution of this technique makes it ideal for studying the real-time processes underlying language comprehension (see Kutas et al., 2006 for a review of language-related ERP research). Effects observed in these studies will provide

useful points of comparison when examining the brain's response to the second element in referential dependencies.

1.2 ERP indices of syntactic dependency formation

The ERP component that would later become known as the phasic² left anterior negativity, or LAN, was first observed in response to morphosyntactic violations in Kutas & Hillyard (1983).³ Phasic LAN effects have since been observed in a range of violation paradigms, but also in response to second elements in well-formed syntactic long-distance dependencies. These include wh-questions in English (Kluender & Kutas, 1993) and German (Felser et al., 2003), relative clauses in English (King & Kutas, 1995; Müller et al., 1997; Weckerly & Kutas, 1999), Japanese (Ueno & Garnsey, 2007), and Korean (Kwon et al., 2013), and in scrambling constructions that disrupt canonical word order in Japanese (Ueno & Kluender, 2003; Hagiwara et al., 2007), Korean (Kwon et al., 2013), and German ((Rösler et al, 1998; Matzke et al., 2002).⁴ Although there are exceptions, whether processing relative clauses, wh-questions, or scrambled word order, the second dependent element typically elicits phasic LAN effects at the post-gap position (though see fn. 4 with regard to LAN responses in head-final relative clauses). Despite this consistency, the functional interpretations proposed for these effects differ, although not necessarily in mutually exclusive ways.

² The term “phasic” is used here to contrast this transient effect (between 300 and 500 or 600 msec. post-stimulus) with longer-lasting sustained anterior negative potentials that often span multiple words.

³ In order to restrict the scope of this review, we will exclude LAN effects elicited in such violation paradigms and other environments that engender “parsing difficulty” (for example Kutas & Hillyard, 1983; Neville et al., 1991; Osterhout & Holcomb, 1992). We note that the functional significance of LAN responses elicited by morphosyntactic violations and those apparently related to verbal working memory processes in grammatical sentences, as discussed in the main text, has never been satisfactorily resolved (though see Martin-Loeches et al., 2005 for a discussion of this issue).

⁴ In most cases, the second element in these dependencies is the gap site, with LAN elicited by the word in post-gap position. However, in head-final SOV languages (e.g. Japanese and Korean) with pre-nominal relative clauses, the head noun is the second element. Even in these cases, LAN-like responses are elicited (see for example Kwon et al., 2013).

1.3 Functional interpretations of phasic LAN effects

Kluender & Kutas (1993) interpreted LAN effects following gap positions in English wh-questions as reflecting the working memory-based⁵ operation of filler retrieval for purposes of filler-gap association. King & Kutas (1995) likewise argued that the phasic LAN effect elicited by the matrix verb in English relative clauses indexes re-activation of the relative clause head noun for purposes of semantic role assignment. Other studies of wh-questions, topicalization (Felser et al., 1997; Matzke et al. 2002), and relative clauses (Weckerly & Kutas, 1999, Ueno & Garnsey, 2010; Kwon et al., 2013) similarly attributed phasic LAN effects to costs related to (some form of) retrieval.

Cross-linguistic studies of scrambling – i.e. leftward displacement of a case-marked noun phrase – have prompted additional interpretations of phasic LAN effects. This is because scrambling often involves nothing more than switching the order of two adjacent noun phrases, making the displacement relatively local (i.e. clause-internal) rather than long-distance (i.e. across clause boundaries). Critical for our purposes is what happens downstream from such scrambled constituents. Ueno & Kluender (2003) observed phasic LAN effects around the object gap position in scrambling constructions, which they interpreted as the parser’s attempt to restore canonical word

⁵ Here we remain largely agnostic on the specific architectural details of the working memory system and its internal operations, relying instead on a simple, more general model. Working memory is a short duration, limited capacity memory system capable of simultaneously storing, manipulating, and combining information in the service of accomplishing some task (Baddeley, 1992). This formulation is sufficient for present purposes, especially in view of general disagreement in the literature about the specific architecture and mechanisms of verbal working memory (see for example Cowan, 1995; Just & Carpenter, 1992; Caplan & Waters, 1999; McElree et al., 2003; Lewis et al., 2006). We will however briefly return to the Lewis et al. (2006) model in the discussion section to aid in the interpretation of results obtained in the current experiment.

order (see Matzke et al. 2002 for similar observations with regard to clause-internal German topicalization).

In Hagiwara et al. (2007), another study of Japanese scrambling, both short- and long-distance scrambling conditions compared to a condition in which constituents were in canonical order elicited phasic LAN effects at the sentence-final main verb. The authors suggested that their sentence-final LAN effects reflected the costs of processing scrambled word order, with concomitant re-computing of relevant semantic roles (cf. King & Kutas, 1995) and grammatical relations.

In sum, the papers cited here agree that phasic LAN indexes a working memory-based operation, though the exact nature of this process appears to be an open question and the inferences made about the underlying processes that LAN effects index rely on assumptions about construction-specific challenges posed to the comprehension system and the precise operations of the working memory system. Accounts based on memory processes such as retrieval, reactivation, (re-)computation and the processing burdens imposed by non-canonicity have been proposed, although many researchers are satisfied with an account that relies on increase in load or cost, broadly construed. Below, we articulate a simplified functional interpretation of phasic LAN that reflects a simple operation of associating two distal elements participating in a long-distance dependency.

1.4 The “back association” hypothesis

Our account of the processes indexed by phasic LAN is not specific to syntactic relationships, does not rely on construction-specific details, nor on specific working memory operations. Rather, on our view, the phasic LAN indexes a process of “back

association” or “association at a distance.” This process is triggered at the second element in a long-distance dependency and consists of the presumed working memory-based operation of “looking back” through already processed material for the first dependent element for the purposes of “associating” it with the second.⁶ This proposal aims to achieve broad empirical coverage by accounting for the totality of LAN effects in syntactic as well as non-syntactic long-distance dependencies.

Perhaps the most appealing aspect of this account of phasic LAN is that it generates testable predictions. Specifically, it predicts that the second of two dependent elements in a sentence-level long-distance dependency will elicit phasic LAN indexing the process by which the second element is associated with the (previously processed) first element. Moreover, this process is not limited to sentence-level syntactic relationships nor to relationships involving null elements (i.e. gaps). If we are correct, then a reassessment of the existing literature should reveal phasic LAN effects beyond those associated with the processing of syntactic filler-gap dependencies, and it is to these effects that we now turn.

1.5 LAN effects to 2nd elements in non-syntactic sentence-level dependencies

Shao & Neville (1998) examined two types of semantic violations that had previously received little attention in the ERP literature. The crucial comparisons were between sentences that violated hyponymy relations (4a) and their controls (4b), as well as between sentences that violated the grammatical requirement for licensing of

⁶ This type of interpretation derives from ideas originally articulated in Kluender & Kutas (1993b) and Kutas et al. (2006).

negative polarity items (NPIs)⁷ (5a) and their controls (5b).

(4a) #Jane does not eat any meat at all, and instead she eats lots of **beef** and vegetables.

(4b) Jane does not eat any meat at all, and instead she eats lots of **rice** and vegetables.

(5a) *Fred believes that he has **ever** seen that woman before.

(5b) Fred believes that he has **never** seen that woman before.

Compared to controls, both violation types (4a & 5a) elicited anterior negativity (albeit with somewhat differing distributions and onset latencies). Shao and Neville relied on a working memory-based account to interpret these effects, noting that processing the hyponymy relation “may require the retrieval of the superordinate (*meat*) when the subordinate (*rice*) is processed,” and that the processing of NPIs may require the parser to “retrieve the initial portion of the sentence to determine whether the negative polarity item fits within its scope” (i.e., within the scope of negation).

Anderson & Holcomb (2005) examined ERP effects associated with the processing of co-reference and synonymy. Participants read two-sentence mini-discourses in which a noun phrase in the second sentence was either intended to be co-referential⁸ (occurring with the definite determiner *the*) or non-co-referential (occurring with the indefinite determiner *a/an*) with a noun phrase in the first sentence. A comparison between co-referential and non-co-referential noun phrases revealed LAN effects in response to nouns following definite determiners compared to those following indefinites. The authors argued that this effect reflected the increased working memory load triggered by the processing of the definite determiner, which signaled that the

⁷ Negative polarity items (such as **ever** in (5a), or **a thing** in “I don’t understand a thing”) are linguistic elements that occur in negative contexts. Compare (5a), in which this requirement is not satisfied for the NPI **ever**, with “Fred **doesn’t** believe that he has **ever** seen that woman before,” in which it is. This contrast illustrates the restrictions on the occurrence of NPIs.

⁸ Put simply, co-referring expressions refer to the same element in either the linguistic discourse or the world.

upcoming noun would require the retrieval of an antecedent in the preceding discourse. These effects, and their interpretations, are consistent with our association-based hypothesis of the functional significance of the LAN, and provide support for our claim that LAN effects may be elicited by the second element in any type of sentence-level relationship, as long as these relationships require the association of two non-adjacent elements.

The present study is explicitly designed to test this hypothesis by investigating ERPs elicited by the second element in simple unambiguous referential dependencies, such as the pronoun *he* in (6):

(6) **Bill** wondered whether **he** would be able to make it to work on time.

Next we summarize what is known about the real-time ERP processing of these referential relationships.

1.6 ERP studies of referential processing

In recent years there has been an increase in the number of ERP studies investigating referential processing (Van Berkum et al., 1999, 2003, 2004, 2007; Anderson & Holcomb, 2005; Heine, 2006; Nieuwland & Van Berkum, 2006; Nieuwland et al., 2007; Osterhout & Mobley, 2007). Van Berkum and colleagues typically have participants read sentences or small discourses, and record ERPs to target pronouns that either have a unique referent (7a), two possible referents (leading to “referential ambiguity,” (7b)), or no referent (leading to “referential failure,” (7c) – examples from Van Berkum et al. (2004)):

(7a) *David* shot at *Linda* as **he** jumped over the fence.

(7b) *David* shot at *John* as **he** jumped over the fence.

(7c) *Anna* shot at *Linda* as **he** jumped over the fence.

Relative to its unambiguous counterpart in (7a), the ambiguous pronoun in (7b) elicited a sustained frontal negativity (approximately between 400 and 1100 msec.), an effect subsequently dubbed the “referentially induced frontal negativity,” or *Nref*. This effect has been elicited using both written and spoken ambiguous pronouns and noun phrases (Nieuwland & Van Berkum, 2006; Van Berkum et al., 1999, 2003). The *Nref* has been argued to reflect situation model-level rather than superficial ambiguity (Nieuwland et al., 2007), correlates with reading span score (Nieuwland & Van Berkum, 2006; Nieuwland et al., 2014), and has been tentatively localized to medial prefrontal cortex (Nieuwland, Petersson & Van Berkum, 2007). *Nref*-like effects have also been reported in response to pro-forms in partially elided noun phrases in ellipsis constructions (Martin et al., 2012; Martin et al., 2014).

Van Berkum and colleagues claim that this response is distinct from the brain’s response to referential failure, observed when comparing pronouns with unique antecedents to those with none, as shown in (7c). The “referentially failing” pronoun in (7c) did not elicit an *Nref* but rather a P600, a late positivity typically associated with ungrammaticality (Hagoort et al., 1993), parsing difficulty (Osterhout & Holcomb, 1992), and long-distance syntactic integration (Kaan et al., 2000). A similar response was reported in Osterhout & Mobley (1995) to anomalous reflexive pronouns, as in:

(8) The man prepared *herself* for the operation.

A few things about this literature are worth noting. None of the manipulations included the straightforward comparison of a pronoun with an overt antecedent to one lacking any kind of candidate antecedent in the discourse, focusing instead on

referential ambiguity or highly salient gender mismatches between pronoun and antecedent. We believe that the latter type of manipulation, which is also a gender-based morpho-syntactic violation, may confound the interpretation of the “referential failure” effects observed. Accordingly, while it has been established that the brain responds differently to different types of referential processing problems, the manner in which the brain processes straightforward referential relationships remains an open question.

To investigate this issue while providing a test of our back association hypothesis, we designed an ERP experiment in which participants read sentences containing pronoun or proper name main clause subjects that either did, or did not, refer back to co-referential elements in a preceding sentence-initial adjunct. In addition, our materials contained an orthogonal manipulation, comparing object relative clauses to their coordinate clause controls, in order to replicate previous findings of phasic LAN in response to the main clause verb in object relative clauses (i.e. King & Kutas, 1995). As such, our study was designed to investigate the ERP signatures of referential dependency formation, while also enabling us to determine the extent to which brain responses elicited by referential dependencies pattern with those elicited by syntactic dependencies.

On the basis of our hypothesis as to the functional identity of phasic LAN, we predicted that at the main clause subject position we would observe phasic LAN effects, i.e. enhanced negativity between 300-500 msec. over left anterior regions of scalp, in response to the two **Co-referent** conditions (conditions A and C, Table 1) compared to their **No Co-referent** (conditions B and D, TABLE 1.1) counterparts.

TABLE 1.1: Experimental sentences: Pronouns and proper names with and without preceding co-referents.

A. Pronoun, Co-referent
After a covert mission that deployed <i>Will</i> for nine terrible months, he longed for home.
B. Pronoun, No Co-referent
After a covert mission that required deployment for nine terrible months, he longed for home
C. Name, Co-referent
After a covert mission that deployed <i>him</i> for nine terrible months, Will longed for home.
D. Name, No Co-referent
After a covert mission that required deployment for nine terrible months, Will longed for home.

We also predicted that the onset and distribution of these negativities would be similar to those elicited by the main clause verb in the comparison between object relative clauses and their coordinate-clause controls (TABLE 1.2) – in other words, we expected the object relatives to elicit enhanced negativity with a similar scalp distribution.

TABLE 1.2: Experimental sentences: Object relative clauses and coordinate clause controls.

E. Object Relative Clauses
The soldier who the sailor roughly pushed smashed a bottle against the bar.
F. Coordinate-clause Controls
The soldier roughly pushed the sailor and smashed a bottle against the bar.

The most commonly held interpretations (Van Berkum et al., 2007) of *Nref* and P600 effects typically elicited in referential paradigms (see section 1.6) would lead one to predict a lack of *Nref* effects in response to our critical comparisons, given the lack of referential ambiguity. With regard to P600 effects, if any pronoun without an obvious intrasentential antecedent is “referentially failing,” the antecedent-less pronoun in the **Pronoun, No Co-referent** condition should elicit a P600 effect. If on the other hand the P600 response to “referentially failing” pronouns is due solely to a morphosyntactic

feature mismatch between a pronoun and a potential antecedent, no such effect should be observed.

2. Results

In this section, we first present the results of the orthogonal comparison between object relatives to their coordinate clause controls in order to establish a point of reference for our primary experimental manipulations, in which we compare the two **Pronoun** conditions and the two **Name** conditions. We do this first for phasic responses in section 2.1, before moving on to more sustained responses in section 2.2.

2.1 Phasic responses

2.1.1 Object relatives versus coordinate clause controls

At the critical verb (*smashed* in E and F of TABLE 1.2), object relatives elicited a negative-going ERP in comparison to coordinate clause controls (FIGURE 1.1).

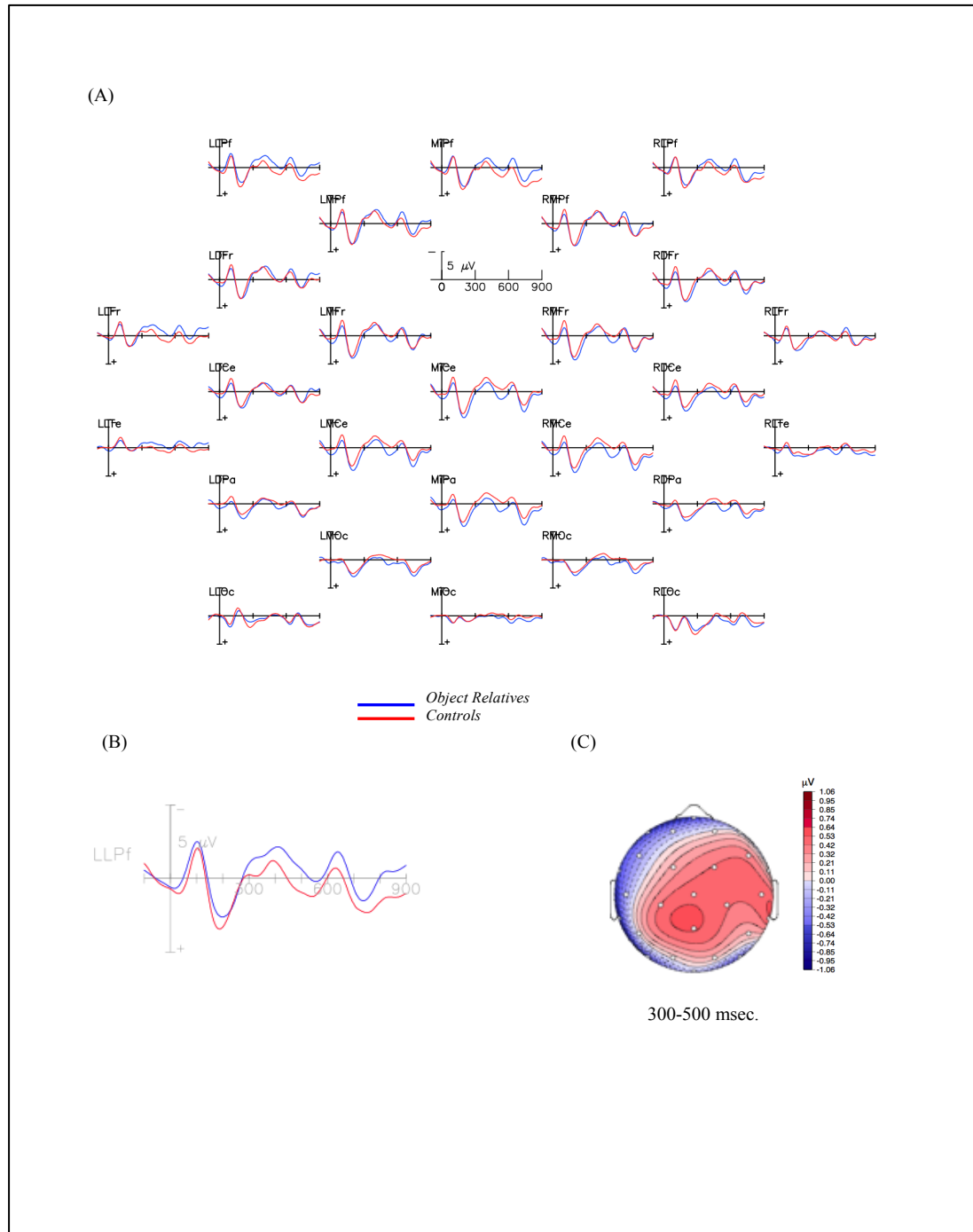


FIGURE 1.1: (A) Grand average ERP waveforms for **object relative** and **coordinate clause controls** at all 26 electrode sites (1000 msec. epoch including 100 msec. baseline). (B) Grand average ERP waveform for object relative and coordinate clause control conditions at electrode LLPf (1000 msec. epoch including 100 msec. baseline). (C) Topographic scalp iso-voltage map of the mean difference between conditions between 300-500 msec.

This effect was small in magnitude and had a narrow distribution, appearing maximal at left antero-lateral scalp sites. A one-tailed t-test on mean amplitude measurements at left anterior electrodes between 300 and 500 msec. confirmed a significant difference between conditions [$t(19) = 2.46$, $p = .02$]. The distributional ANOVA revealed no main effect of condition [$F < 1.0$], but there were significant condition x hemisphere [$F(1,19) = 6.16$, $p = .02$], condition x laterality [$F(1,19) = 5.85$, $p = .0384$], and condition x hemisphere x laterality [$F(3,57) = 8.57$, $p = .015$] interactions, as well as a marginal interaction of condition x anteriority x hemisphere x laterality [$F(3,57) = 2.63$, $p = .08$]. These interactions appeared to be driven by a difference between conditions at left lateral scalp sites [$F(1,19) = 9.82$, $p = .006$] that was not present at left medial, right medial, or right lateral sites (all $F_s < 1.0$). Additionally, though the marginal four-way interaction suggested that the differences between conditions might vary at different levels of the anteriority factor, follow-up analyses showed this not to be the case, as comparing at individual levels of the anteriority factor over the left hemisphere yielded no significant results (all $F_s < 2.0$).

In sum, the comparison between object relatives and their coordinate clause controls (E vs. F in Table 1.2) yielded LAN in response to the critical main verb in the object relative clause condition. This LAN effect, which exhibited a left lateral distribution (see Figure 1.1 and section 3.3), was significant in the traditional LAN time window of 300-500 msec., replicating previous findings in the literature. This effect at the second element of a syntactic dependency can now serve as a point of comparison for the brain response to the second element in our referential dependencies.

2.1.2 Experimental sentences

The results reported in this section focus on responses occurring at least 300 msec. post-onset of the critical main clause subject. No comparisons undertaken in the N100 (50-150 msec.) and P200 (150-250 msec.) latency windows at this sentence position produced significant results.

2.1.2.1 Pronouns with and without co-referents

2.1.2.1.1 Responses between 300 and 500 msec.

Visual inspection of the ERP response to the pronominal subject suggested that the response to the **Pronoun, Co-referent** condition was more negative than the response to its **No Co-referent** counterpart, as shown in FIGURE 1.2.

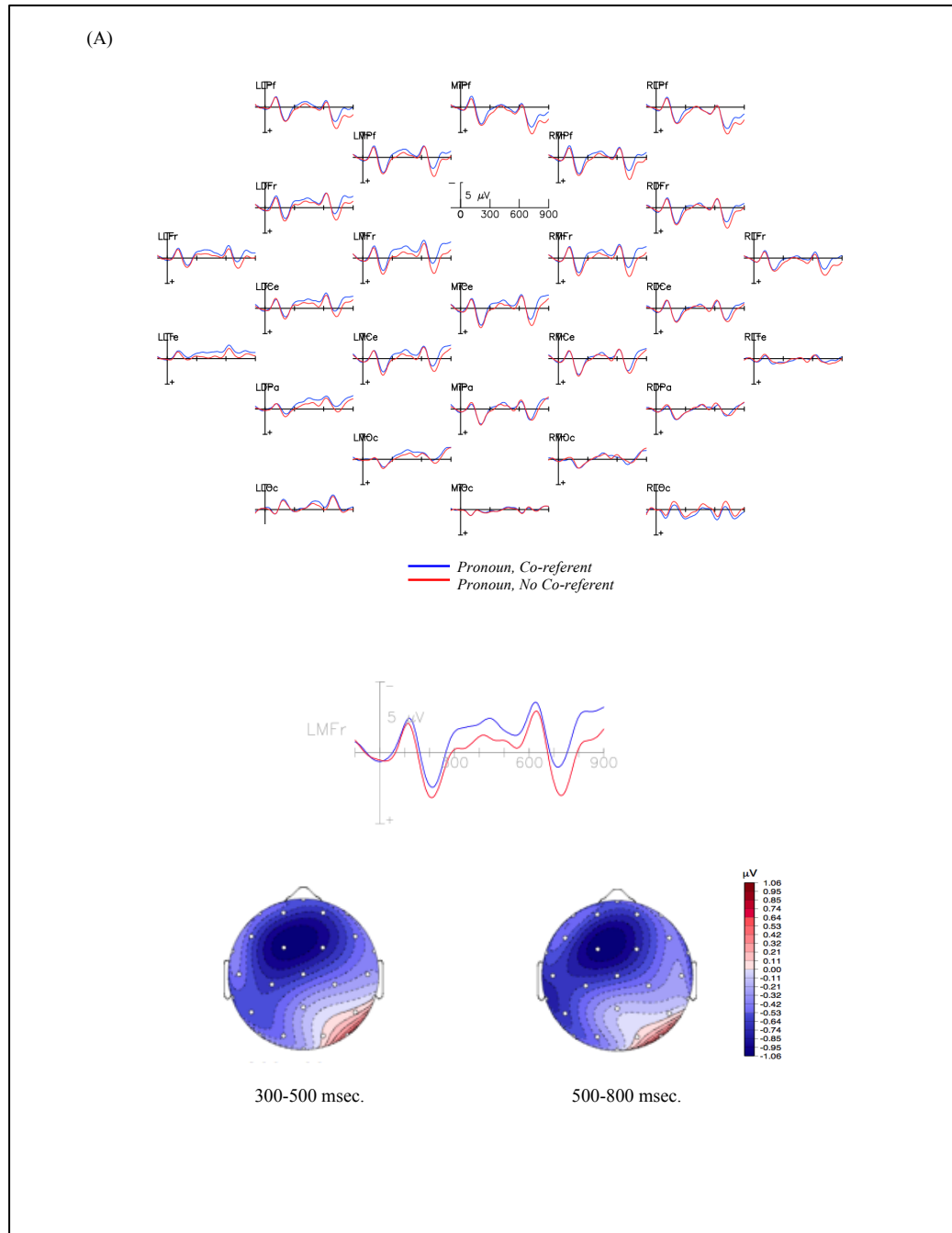


FIGURE 1.2: (A) Grand average ERP waveforms for the **Pronoun, Co-referent** and **Pronoun, No Co-referent** conditions at all 26 electrode sites (1000 msec. epoch including 100 msec. baseline). (B) Grand average ERP waveforms for both **Pronoun** conditions at electrode LMFr (1000 msec. epoch including 100 msec. baseline). (C) Topographic scalp iso-voltage map of the mean difference between conditions in the 300-500 msec. time window. (D) Topographic scalp iso-voltage map of the mean difference between conditions in the 500 - 800 msec. time window.

This effect appeared larger over the left than the right hemisphere, and also appeared to be slightly larger at anterior electrodes. A one-tailed t-test on mean amplitude measurements at left anterior electrodes between 300 and 500 msec. confirmed a significant difference between conditions [$t(19) = 4.42, p < .0005$]. The distributional analysis showed no main effect of condition [$F < 1.0$], but did reveal a significant condition x hemisphere interaction [$F(1,19) = 5.78, p = .0049$], driven by a main effect of condition over the left [$F(1,19) = 5.61, p = .02$] but not the right hemisphere [$F < 1.0$]. There were also marginal condition x laterality [$F(1,19) = 4.14, p = .0577$] and condition x hemisphere x anteriority [$F(3,57) = 2.22, p = .09$] interactions in the distributional analysis. Investigating these interactions at individual levels of these factors, there were main effects of condition at left frontal [$F(1,19) = 4.13, p = .05$] and left temporal [$F(1,19) = 8.51, p = .0096$] sites, but not at left prefrontal [$F(1,19) = 1.07, n.s.$], left occipital [$F(1,19) = 2.74, n.s.$], or right hemisphere sites (all $F_s < 1.0$), and that effects were significant at left medial [$F(1,19) = 6.94, p = .024$], but not left lateral, right medial or right lateral sites (all $p > .27$). The isovoltage map in Figure 2.1C shows the left fronto-central distribution indicated by the results of these comparisons.

2.1.2.1.2 Responses between 500 and 800 msec.

Visual inspection of the data also suggested increased positivity in response to the **Pronoun, No Co-referent** condition, as predicted by the work of Van Berkum and colleagues. Results of planned t-tests confirmed significant differences between conditions. In the 500 to 800 msec. time window, the distributional analysis correspondingly revealed a marginal main effect of condition [$F(1,19) = 3.99, p = .067$],

as well as a significant condition x hemisphere interaction [$F(1,19) = 4.85, p = .0417$] and a marginal condition x hemisphere x anteriority interaction [$F(3,57) = 2.10, p = .10$] – in other words, the same interactions with distributional factors as were reported for the 300-500 msec. time window, minus the marginal interaction of condition x laterality. Perhaps unsurprisingly, the isovoltage map in FIGURE 1.2D indicates that the distribution of the response in the later time window of 500-800 msec. was virtually identical to the distribution of the LAN effect observed in the earlier 300-500 msec. window. We therefore provisionally conclude that what might initially appear to be a late positivity in response to the **Pronoun, No Co-referent** condition is in fact another phasic LAN response to the word subsequent to the critical main clause subject position. We return to this issue in section 3.1.1.

2.1.2.1.3 Summary

In sum, while the onset latency, polarity, and lateralization of the phasic LAN effects elicited by both object relatives and pronouns with antecedents were similar as predicted, the effects did differ subtly in terms of their distribution across the scalp. The phasic LAN effect elicited by object relative clauses was significant at left lateral scalp sites (with no difference between conditions at left medial scalp sites), but did not differ reliably along the anterior-posterior dimension. In contrast, the LAN effect elicited by pronouns with antecedents in the **Pronoun** comparison *did* differ along the anterior-posterior dimension, but did not differ in comparisons across conditions at left medial or left lateral electrodes. Last but not least, based on the scalp distribution of the potential observed in the 500-800 msec. time window, we interpret this response as a phasic LAN

response to the word one position downstream from the critical position rather than a late positivity.

2.1.2.2 Proper name subjects with and without co-referents

In comparing the **Name, Co-referent** and **Name, No Co-referent** conditions (conditions C and D, TABLE 1.1) at the matrix clause subject position, not only were there no effects in the N100 and P200 time windows, but additionally, neither the results of planned t-tests nor the full or distributional analyses in any of the later time windows (i.e 300-500 or 500-800 msec.) tested showed main effects of condition or any interactions with electrode or distributional factors (see FIGURE 1.3).

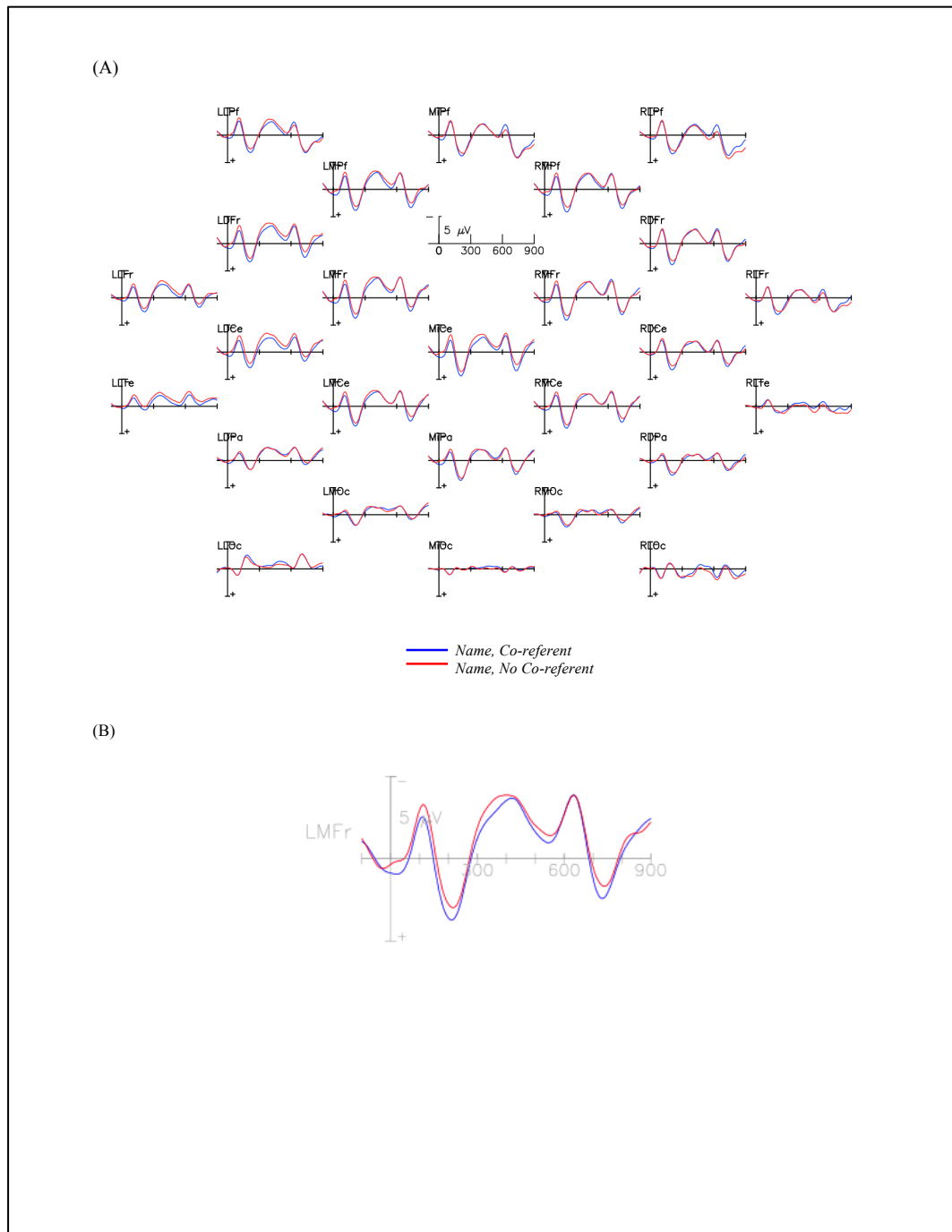


FIGURE 1.3: (A) Grand average ERP waveforms for the **Name, Co-referent** and **Name, No Co-referent** conditions at all 26 electrode sites (1000 msec. epoch including 100 msec. baseline). (B) Grand average ERP waveforms for both **Name** conditions at channel LMFr (1000 msec epoch including 100 msec baseline). No iso-voltage map is provided because of a lack of observable differences between conditions.

2.2 Sustained responses (300 – 2000 msec)

2.2.1 Object relatives versus coordinate clause controls

As has previously been reported for LAN effects elicited by second elements in English object relative clauses (King & Kutas, 1995) and wh-questions (Kluender & Kutas 1993a), the response to the main clause verb in our orthogonal manipulation appeared to continue past the critical word (see Figure 4.1A). The full multi-word⁹ analysis conducted between 300 and 2000 msec. (including the critical word plus three additional words: *smashed a bottle against* in E and F of Table 1.2) revealed a significant condition x electrode interaction [$F(25,475) = 2.22$, $p = .0007$], and the distributional analysis showed a marginal condition x anteriority interaction [$F(3,57) = 2.48$, $p = .0718$] as well as a significant condition x anteriority x hemisphere interaction [$F(3,57) = 4.41$, $p = .0078$]. Further investigation of these interactions indicated that the difference between object relatives and controls was larger over the left hemisphere, and slightly larger over frontal sites. However, the individual word analyses of the three words subsequent to the critical position revealed no significant effects after re-baselining (all $F_s < 1.0$; see isovoltage maps in Figure 4.1B).

2.2.2 Pronouns with and without co-referents (300-2000 msec.)

Similar to the LAN effect elicited by object relatives vs. coordinate clause controls, the LAN effect elicited by the main clause pronominal subject with a preceding co-referent also seemed to extend over multiple word positions.

The full multi-word analysis between 300 and 2000 msec. (including the entire main clause of conditions A and B in Table 1.1: *he longed for home*) again showed a

⁹ See section 4.5 for a discussion of the distinction between single- and multi-word analyses of sustained effects.

significant condition x electrode interaction [$F(25,475) = 1.98, p = .037$], and the distributional multi-word analysis again showed a significant condition x anteriority x hemisphere interaction [$F(3,57) = 3.69, p = .0177$] attributable to large differences between conditions over the front of the head, predominantly over the left hemisphere (see Figure 1.5A). In this respect, the sustained effects elicited by pronouns with antecedents were very similar to those elicited by the main clause verbs in object relative clauses.

However, in contrast, individual word analyses of the sustained effect elicited by pronouns with antecedents (Figure 1.5B) showed a significant main effect of condition at word positions 13 (*longed* in conditions A and B of Table 1) [$F(1,19) = 3.89, p = .04$] and 14 (*for* in conditions A and B of Table 1) [$F(1,19) = 4.57, p = .0473$], as well as condition x laterality [$F(1,19) = 6.14, p = .02$] and a marginal condition x laterality x anteriority [$F(3,57) = 3.94, p = .11$] interactions at word position 14. These interactions were caused by differences between conditions mainly over fronto-central scalp sites.

3. Discussion

Generally, the results of this study provide evidence in support of our back association hypothesis. Consistent with our predictions, we observed greater left anterior negativity in response to pronouns with preceding antecedents compared to those without, arguably indexing the process of associating a pronoun with its antecedent, a process necessary for the establishment of a straightforward referential dependency, much like the syntactic dependency between a filler and its gap. Somewhat to our surprise, we observed no such effect in our **Name** comparison, an issue to which we return below in section 4.2.

The similar (albeit non-identical) within-participant effects observed at the second element in both syntactic and referential dependencies – at least as far as simple antecedent-pronoun relations are concerned – suggest that phasic LAN-like effects reflect a cognitive process that is largely independent of the linguistic level of analysis at which it occurs. However, subtle differences in the duration of effects and scalp topography suggest that back association processes and their attendant brain responses may be modulated to some extent by the type of long-distance relationship under construction. We return to this question in section 4.3.

3.1 Pronominal main clause subjects (300-500 msec.)

Based on our hypothesis regarding the functional identity of LAN effects, we predicted enhanced LAN in response to a pronoun preceded by a co-referring nominal antecedent in the same sentence when compared to an antecedent-less pronoun, and this is the pattern of results that we observed. Accordingly, we maintain that the LAN effect observed in the comparison of the **Pronoun** conditions can plausibly be interpreted as an index of the association processes necessary for dependency formation (FIGURE 1.2A & 1.2C).

An alternative, albeit less likely, interpretation of this LAN-like negativity in the current study *is that it is* an *Nref* indexing referential ambiguity, an interpretation buttressed by the results of an analysis in the traditional *Nref* time window of 400 - 1100 msec. that showed significant differences between conditions (condition x anteriority x hemisphere interaction [$F(3,57) = 4.17, p = .0113$]). If this were the case, however, one would be forced to adopt the position that the parser diagnoses the main clause subject in the **Pronoun, Co-referent** condition as referentially ambiguous, with the ambiguity

(perhaps) arising because of indecision as to whether to associate the pronoun with (a) the co-referring nominal expression in the preceding sentence-initial adjunct or (b) an extra-linguistic referent. However, post-experiment debriefing questionnaires provided no evidence that participants experienced these items as ambiguous, or even difficult to process.¹⁰ Moreover, during debriefing, no participants indicated that they had been associating pronouns with extra-linguistic co-referents. Nor do we believe that this would have been a reasonable default parsing strategy, given the presence of an explicit co-referent with matching number and gender features within the same sentence.

If we pursue the line of reasoning that the enhanced negativity in response to the pronoun is an index of referential ambiguity (i.e. an *Nref*), we are inevitably led to the conclusion that the parser initially experiences *every* non-bound-variable pronoun (Reinhart, 1983) as ambiguous – a position that does not seem parsimonious or consistent with evidence that the processor often initially pursues the simplest parse available (Frazier, 1987). We therefore maintain that the negativity observed in response to the pronominal subject in the **Pronoun, Co-referent** condition is a LAN most plausibly interpreted as indexing the process of associating two non-adjacent elements.

3.1.1 Pronominal main clause subjects

While visual inspection of the ERPs to **Pronouns** with and without co-referring antecedents (FIGURE 1.2) might suggest a relative positivity in response to the **No Co-referent condition**, careful examination of the data renders this interpretation unlikely.

¹⁰ In fact, during debriefing, the majority of subjects reported having the most difficulty processing object relative clauses, assuming that it was these constructions that were the focus of the experiment.

We instead interpret this effect as a LAN in response to the word subsequent to the critical position. Support for this view comes from the virtually identical scalp topographies of this comparison in these two adjacent time windows (see the isovoltage maps in FIGURES 1.2C & 1.2D), the fact that our analyses showed, after re-baselining, a significant LAN response to the subsequent word, and the fact that effects observed in the 500-800 msec. time window were preceded by an epoch (400-600 msec.) during which the comparison between the two pronoun conditions did not reach significance ($p > .35$). Moreover, our late effect does not resemble the late positivities observed in response to antecedent-less pronouns in the work of Van Berkum and colleagues, which all present with a clear posterior maximum (Van Berkum et al., 1999, 2003; Nieuwland & Van Berkum, 2006).¹¹

The apparent absence of a late positivity to the **Pronoun, No co-referent** condition calls into question the assertion that pronouns lacking antecedents universally elicit late positivity. As mentioned in the introduction, in the most widely used manipulation intended to induce referential failure, co-referents are typically rendered unavailable with a gender or number manipulation that results in a highly salient feature-based mismatch between pronouns and their practically adjacent (potential) antecedents. We suggest that it is this mismatch that leads to the late positive response – arguably a P3b (Donchin, 1981; Kok, 2001) or a P600 effect, elicited not by a

¹¹ It should also be noted that a previous study investigating the introduction of new referents to an existing discourse (Kaan et al., 2007), also elicited late positivity. However, in line with the argument above, this effect had a centro-lateral distribution and was significant between 900-1400 msec., providing further support that we are not observing frontal positivity in response to our **Pronoun, Co-referent** condition, as we can find no precedent for such an effect in the literature.

referentially failing pronoun, but rather by a salient gender-based morphosyntactic anomaly (Osterhout & Mobley, 1995; Coulson et al., 2001).

As also noted in the introduction, the ERP literature on referential processing contains many manipulations that, though they may not always create outright violations *per se*, at a minimum engender processing difficulty as the parser attempts to construct referential relationships. The most frequently cited response to such manipulations is the *Nref*, a frontal negativity linked to processing costs associated with referential ambiguity. In our study, we elicited negativity that was very similar in timing (see section 4.3.1) and scalp distribution (see section 4.3.2) to the *Nref* in response to the second element in an unambiguous, straightforward referential dependency..This suggests that the *Nref* may not be a specific response to referential ambiguity alone. Thus while we agree that the *Nref* indexes referential processing, we question whether the presence of referential ambiguity is a necessary condition for its elicitation. Instead, we would like to suggest that the *Nref* is just one instance of a family of anterior negativities elicited by association at a distance in language contexts more generally (this conjecture is not very different from one found in Van Berkum et al., 2007),.

We believe that the results of our study can be reconciled with the totality of findings reported across studies by Van Berkum and colleagues with a slight modification of the cognitive processes that the *Nref* is assumed to index. Specifically, the LAN response elicited by pronouns with antecedents in our materials was found to be larger than that elicited by pronouns without antecedents. Similarly, Van Berkum and colleagues have shown that the *Nref*, with very similar timing parameters and scalp topography to our data, is larger in response to pronouns (and other anaphoric

expressions) with two possible antecedents than to pronouns with only one possible antecedent. Under this view, LAN effects in response to pronouns with antecedents are interpreted as indexing straightforward dependency formation, while *Nref* effects elicited in response to pronouns with multiple candidate antecedents index *attempts* to form such a dependency, with increases in amplitude the result of the difficulty associated with selecting among multiple candidate antecedents.

In other words, the amplitude of the anterior negative response in each case indexes association at a distance of two dependent sentence elements, with the amplitude of the response determined by the ease of association. In line with our study showing similarities between syntactic and referential dependencies, this view is consistent with data reported in King & Kutas, (1995:389, Figure 6) showing monotonic increases in the amplitude of LAN responses from coordinate (no dependency) to subject relative (dependency between main clause subject and verb) to object relative clauses (dependency between main clause subject and verb, and between filler and gap). If this interpretation is accurate, it would predict that a pronoun with three possible antecedents might likewise elicit greater anterior negativity than one with two, which is an empirical question for future research.

3.2 Proper name main clause subjects

Counter to our expectations, we observed no differences between the two **Name** conditions at the main clause subject position. This leads us to conclude that – unlike in the **Pronoun** conditions – back association operations were not triggered by the proper names in main clause subject position of our stimulus sentences.

This result would seem to be in direct conflict with data showing an increase in reading time at the second element (i.e. a proper name) in constructions in which the pronoun precedes its co-referent (Kazanina et al., 2007). This increase is argued to reflect processes underlying the formation of the referential dependency between a pronoun and a subsequent co-referring proper name. However, Kazanina et al. argue (2007:406) that if the parser encounters evidence (in our case the initial adverbial adjunct) predicting an upcoming clause that will contain a subject (as was always the case in our materials), this information could influence the parse before the subject is actually encountered. If this is the case, then there is no reason to assume that the proper name subjects would necessarily trigger back association operations.¹²

Put simply, under our hypothesis, if the process of back association is not triggered, then no LAN effect should be elicited. This account gains support from both theoretical and experimental work on reference form, as well as from recent influential models of the working memory processes subserving sentence comprehension (McElree et al., 2003; Lewis et al., 2006).

It has been noted (Ariel, 1990; Gundel et al., 1993) that all known languages contain a range of referential forms that vary in terms of lexical specificity/elaboration. For example, in English, these range from null forms to proper names, with several additional intermediate forms. It has been argued that these referential forms constitute

¹² In fact, as Kazanina et al. argue, it is probably the processing of the first element in the dependency, namely the pronoun, which imposes the processing burden on the parser. Unfortunately, as discussed below (see fn. 20), there was despite our best intentions no suitable comparison for this pronoun preceding its proper name co-referent in our materials, and it was thus impossible for us to determine the electrophysiological indices of the processing burden imposed by these constructions. It seems reasonable to us to suppose that such an electrophysiological response may have been present in our data, but this issue will need to be resolved in future research.

a hierarchy, with fully specified forms at one end of the spectrum and attenuated (or even elided or implicit) forms at the other, with position on the hierarchy argued to predict, most importantly for our purposes, *referential role*. Crucial for present purposes is the contrast between pronouns and proper names, elements that occupy opposite poles of the referential hierarchy and that are argued to fulfill the roles of *referential maintenance* and *referential establishment*, respectively (Silverstein, 1976). If we adopt this logic, then we can straightforwardly argue that while pronouns typically “look backward” to maintain reference, proper names function to establish reference and intrinsically “look forward.” This notion gains support from experimental work by Marslen-Wilson et al. (1982) and Vonk et al. (1992) that demonstrates the different roles these forms play: (1) pronouns are associated with the continuation of existing topics and the maintenance of focus on antecedents, while (2) proper names tend to introduce new topics and referents. We therefore believe that the referential hierarchy provides a simple explanation of the null effect in the **Name** condition comparison: we observed no LAN effect in the **Name Co-referent** vs. **Name, No Co-referent** comparison because no back association operations were initiated.

In addition to this work on reference form, the model in Lewis et al. (2006), based on well-established memory constructs such as interference, encoding, and retrieval, is intended to explain numerous effects in the sentence processing literature (though as stated above, here we do not commit to a specifically retrieval-based functional interpretation of observed LAN effects). Most pertinent to our purposes is the notion of cue-based content-addressable retrieval (a notion first elaborated in McElree et al., 2003). Simplifying for ease of exposition, features of certain words in

the language input trigger the retrieval of previously processed items in order to associate them. The retrieval operation is “content addressable,” (i.e., access is direct rather than serial), and cue-based, such that the retrieval is cued by features, and retrieval efficiency is dependent on the number of items in memory that share features with the cue.

Under this set of assumptions, it becomes easy to explain the contrast between our **Pronoun** and **Name** comparisons. The second element in the **Pronoun, Co-referent** condition is a pronoun, which could plausibly be assumed to have a feature that cues the retrieval of a suitable antecedent (presumably in addition to, for example, number and gender cues), particularly when a candidate antecedent is available in the same sentence. In contrast, no such [+antecedent] feature is associated with the second element (a proper name) in the **Name Co-referent** condition, and therefore retrieval can plausibly be assumed to not be cued.

Additionally, we believe that the model in Lewis et al. (2006) provides further support for our assertion that the LAN effect in response to our **Pronoun, Co-referent** does not index the mechanisms underlying the processing of referential ambiguity, but rather a negativity that is part of a continuum, the amplitude of which is determined by the difficulty of forming a referential dependency. Specifically, there is no need for the parser to search for an antecedent outside the current discourse in our experimental paradigm (but see Nieuwland (2014) for an experimental paradigm very different from ours in which participants appear to be doing just that), given that there is a readily available, featurally appropriate, and syntactically licensed antecedent within the current sentence. In sum, this architecture helps to explain our pattern of results, and in

combination with the work on reference form, supports our contention that no back association operation is triggered in the **Name, Co-referent** condition.

3.3 Relation between syntactic and referential dependencies

An important part of our argument for a domain-general back association interpretation of the LAN rests on its elicitation by the process of association in the comparison between object relatives and their controls, as well as in the comparison between the two **Pronoun** conditions. In support of this view, the second element in both syntactic and referential dependencies elicited LAN in our data. At the same time, these negativities were not identical in either their extended time course or in their amplitude distributions across the scalp, even within the same group of participants. Although this may seem non-optimal for our position, it is also not sufficient to undermine our argument, as discussed in the following sections.

3.3.1 The time course of anterior negativity

Brain responses to the second elements in syntactic dependencies (i.e. the main clause verb following the gap position at the end of the relative clause; TABLE 1.2, condition E) and referential dependencies (i.e. the main clause pronoun subject following the sentence-initial adjunct; TABLE 1.1, condition A) appeared to persist beyond the critical word, as confirmed by interactions of these LAN effects with factors of anteriority and hemisphere in the multi-word analyses (see section 2.5) between 300 and 2000 msec. post-onset of the critical word (section 3.2 and FIGURES 1.4A & 1.5A). The responses to syntactic and referential dependencies were thus very similar in eliciting apparently sustained effects.

However, the responses to syntactic and referential dependencies differed in the individual word analyses undertaken between 300 and 500 msec. post-onset of the words following the critical word. In the individual word analyses of the object relative vs. conjoined clause comparison, there were no further significant differences after the critical word when the onsets of subsequent words were rebaselined. In other words, only the critical word, the main clause verb in object relative clauses, elicited a LAN effect in this comparison; subsequent words apparently sustained the effect but did not contribute to it independently. On the other hand, the individual word analyses of pronouns with and without antecedents showed a significant difference between 300 and 500 msec. not only in response to the critical word, but also in response to the two words following it, even under rebaselining. This suggests that the two words following the pronoun subject in this comparison made independent contributions to the LAN effect.

However, we believe that the response to the critical word in the pronoun comparison was entirely independent of the responses to the two following words, for the following reason: when we measured 400 to 600 msec. post-onset of the critical word (i.e. the last 100 msec. of the response to the critical word and the first 100 msec. of the response to the following word), we found no significant differences. This suggests that the LAN effect in response to the critical word did not spill over into the following word (*longed* in TABLE 1.1), but rather that this word appears to have elicited an independent LAN effect of its own. We similarly found no significant differences in an analysis of the 400 to 600 msec. time window post-onset of word 13, suggesting that the LAN effect at this position was also independent of the LAN effect

elicited by word 14 (*for* in TABLE 1.1). This difference in the individual word analyses suggests that incoming words following the second element in a referential dependency induce additional processing difficulty, while those following the second element in a syntactic dependency do not. Why might this be the case?

While we do not wish to make strong commitments to the nature of the underlying cognitive processes involved on the basis of the data from this study alone, we nonetheless note that the prolonged LAN effects in response to pronouns with antecedents in our materials appear consistent with the serial processes described in a two-stage model of real-time referential processing proposed by Garrod & Sanford (1990). During the initial automatic bonding stage, the comprehension system forms a loose and superficial attachment between a pronoun and possible antecedents, a process that can tentatively be mapped to the back association process that we argue is indexed by phasic LAN. A full referential interpretation of the pronoun is only established in the subsequent resolution stage. The resolution process itself can be temporally prolonged, especially in cases in which there is no focused antecedent in the preceding discourse. In these cases, resolution may not be completed until disambiguating information downstream from the pronoun becomes sufficient to establish a one-to-one mapping, and it may be this resolution process that is indexed by the effects observed at the words downstream of the matrix clause subject pronoun in condition A of our materials. In any case, the differences in time course elicited during the processing of referential and syntactic dependencies point to similar but non-identical mechanisms underlying the processing of these long-distance relationships.

3.3.2 *The scalp topography of anterior negativities*

Differences in the scalp topography of anterior negativities, both phasic and sustained, in response to syntactic dependencies are not without precedent in the literature: though often reported with canonical left anterior maxima (see introduction), bilateral anterior (King & Kutas, 1995; Ueno & Kluender, 2003) and right anterior (Müller, King & Kutas, 1997; Ueno & Kluender, 2009; Dillon et al., 2012) distributions are also common. Moreover, although one cannot infer the location of neural generators for ERP effects from their amplitude distribution at the scalp, the fMRI literature suggests reasonable, overlapping but non-identical neural sources for the negativities elicited in response to the second element in a long-distance dependency. Recall (section 1.6) that Nieuwland et al. (2007) tentatively localized the generators of *Nref* effects in response to referential ambiguity to medial prefrontal cortex. In contrast, an array of fMRI and PET studies have typically reported increased activation in response to long-distance syntactic dependencies in traditional language areas of left lateral cortex. For example, Just, Carpenter, & Keller (1996) compared object relative clauses, subject relatives, and coordinate clause controls (such as were used in our experiment) and reported the greatest activation in and around Broca's and Wernicke's areas of the left hemisphere. Other neuro-imaging studies (Stromswold et al., 1996; Caplan, Alpert, & Waters, 1998, 1999; Caplan et al., 2000, 2001; Cooke et al., 2001; Ben-Shachar and colleagues, 2003, 2004) have likewise identified anterior and posterior (though not always both) language-related areas of activation on the lateral surface of the left hemisphere in response to the more complex of two constructions, in most cases when comparing object to subject dependencies. Crucially for our purposes, none of the

reported loci of activation were in medial frontal areas. Though these inferences remain somewhat speculative, these neuro-imaging studies offer a plausible account of the differences in the scalp distributions of the negativities observed in response to the second element in referential and syntactic dependencies, which thus far appear to be have medial and lateralized generators, respectively.

It is true that that differences in both the time course and the scalp distribution of the effects elicited by the second element in our syntactic and referential dependencies could turn out to be specific to the group of participants we tested and/or the set of materials we used,¹³ and therefore that the underlying mechanisms involved in processing these dependencies are more similar than our data suggest. After all, contra to our findings, previous studies investigating syntactic object dependencies have reported LAN effects extending beyond critical words following gapped positions, even after re-baselining (Kluender & Kutas 1993a:205-6; King & Kutas, 1995:386-7). Likewise, it is possible that the somewhat atypical left-lateral scalp topography of the LAN response to object relative sentences in our results¹⁴ is specific to the particular experimental circumstances of our study. However, even if this should turn out to be the case, either in terms of time course or of scalp distribution, it would only lend further credence to our back association hypothesis, which predicted LAN effects in response

¹³ Somewhat atypically, our materials did not include subject relative clauses as a point of comparison for object relative clauses, but rather coordinate clause controls matched to the relative clause sentences in lexical content (which was not the case for example in the materials used by King & Kutas, 19995). Additionally, our object relative clauses were by design and intent embedded in a set of sentences containing various types of referential relationships, which obviously provides a very different context from those used in previous studies examining the processing of syntactic dependencies.

¹⁴ In fact, subsequent unpublished work conducted in our lab comparing these same object relative and coordinate clause sentences elicited more standard left anterior responses when not inter-mixed with conditions containing referential dependencies.

to the second element in both dependency types. If future work using different materials shows the time course and distribution of these effects to be even more similar than those elicited in our study, this would still be entirely consistent with our claims.

As stated in the introduction, there has been a longstanding debate within the theoretical linguistics literature (Chomsky, 1980, 1981; Reinhart & Reuland, 1993) as to exactly which types of referential relationships are handled within the syntax and which are handled via pragmatics and principles of discourse. While it is too early to make any concrete claims based on electrophysiological evidence, we believe that experiments like the one reported here represent a potentially fruitful way to tease these issues apart and begin to understand the cognitive processes involved in the formation of various types of long-distance relationships.

4. Methods and Materials

4.1 Participants

20 monolingual native speakers of English participated in the experiment (9 females, 11 males; mean age = 20.7), and received either course credit or \$7 an hour for their participation. All students were enrolled as students at the University of California, San Diego. All participants were right-handed with no neurological or psychiatric disorders and normal or corrected-to-normal vision.

4.2 Materials

A total of 160 sets of the four conditions shown in TABLE 1.1 were created for the purposes of this experiment. Each experimental item was fifteen words long, and consisted of an eleven-word sentence-initial adjunct followed by a four-word main clause. Co-referents of the main clause subject occupied the same position in the

sentence-initial adjunct across the two **Co-referent** conditions (condition A: co-referential antecedent, condition C: preceding co-referential pronoun), and the four sentence positions intervening between the first and second co-referents were identical in the main comparisons, establishing a clean baseline for our critical comparisons, which were made at the subject position introducing the main clause, and in all cases consisted of comparisons between identical lexical items.

The sentence-initial adjuncts in the primary experimental conditions all had either *after*, *because*, *since*, *until*, or *during* in initial position, with the occurrence of these adverbials balanced across items (32 items per initial adverbial). The verbs internal to these adjuncts, crucial to the experimental manipulation, were selected in one of two ways. First, as in the example in TABLE 1.1, we selected transitive verbs that, when nominalized, could serve as the internal argument of verbs like *require* or *provide*. In this manipulation, co-referents served as direct objects of the verb (e.g. *deployed Will/him*) in the two **Co-referent** conditions, while the nominalization eliminated the co-referent (e.g. *required deployment*) in the two **No Co-referent** conditions. Second, verbs were selected from the list of “*pro-arb*” verbs in Levin (1993). These verbs, which select for optional direct object arguments, were employed in their transitive forms in the two **Co-referent** conditions (e.g. *startled Katherine*) and were changed into an adjectival predicate form in the two **No Co-referent** conditions (e.g. *was startling*). These verb selection strategies ensured that there was minimal repetition of adjunct-internal verbs across items, with a single verb repeated no more than four times across experimental items. The object of the sentence-initial preposition, an adjective-noun combination like *covert mission* in TABLE 1.1 (which also served as

the head noun of the subject relative clause internal to the sentence-initial adjunct) was unique in each item set and held constant across conditions.

80 additional sentences constituted an orthogonal manipulation of crucial interest, namely 40 object relative clauses (condition E, TABLE 1.2) and their coordinate clause controls (condition F, TABLE 1.2); cf. King & Kutas (1995).¹⁵ By including these syntactic dependencies in the materials, we were able to look for similarities and differences in the brain responses elicited by these dependencies compared to those elicited by referential dependencies in the same group of participants. The set of materials also contained 80 filler sentences, shown in Table 1.3. In an attempt to provide a point of comparison for the adjunct-internal pronoun in the **Name, Co-referent** condition C, subjects read 40 fillers in which an adjunct-internal pronoun referred back to a preceding sentence-initial antecedent (condition G, Table 3). In retrospect, this comparison turned out to be confounded and will not be discussed further.¹⁶ 40 of the other fillers were similar to the experimental sentences, and were designed to add variety to the materials. 30 of these 40 contained sentence-medial (rather than sentence-initial) adverbials (*after, because, since, until, during, and, in addition, before*) and varied in terms of both number and type (common versus proper) of nouns present (condition G, TABLE 1.3). The remaining ten filler items were

¹⁶ In constructing Condition G as a control for Condition C, we failed to take into account the fact that these two conditions might elicit responses for different reasons – one has a preceding antecedent, much like the main clause subject comparison, while condition C contains a pronoun preceding its antecedent, also known as a cataphor. Thus the comparison was confounded.

identical in structure to the experimental items, except that the sentence-initial adjunct began with *before*¹⁷ (condition H, TABLE 1.3).

TABLE 1.3: Filler sentences

G. Adjunct-internal Anaphor with Sentence-initial Antecedent
<i>Will</i> joined a mission that deployed him for nine terrible months and longed for home.
H. Sentence-medial Adverbials
Lori nailed the interview after Lawrence coached her on how to impress the recruiter.
I. Before-Adjuncts
Before a covert mission that deployed <i>him</i> for nine terrible months, Will longed for home.

The resulting set of 320 sentences was rotated through a Latin square design into ten lists, with two orders per list, for a total of 20 lists. The sentences in each list were pseudo-randomized such that three sentences of the same condition never appeared consecutively, and there were never more than five items from the same condition within a sequence of ten items.

4.3 Procedures

Subjects were run in a single session that lasted approximately 2.5 hours, including preparation. The experiment consisted of five blocks, each containing 64 sentences and lasting approximately 15-17 minutes. During the session, subjects were seated comfortably in a chair in a sound-attenuated booth. 1500 msec. before the first word of a sentence, a red square appeared in the middle of the screen and remained throughout sentence presentation for fixation purposes. Sentences were visually presented above fixation, with each word presented for 300 msec. (500 msec. stimulus onset asynchrony (SOA)). True/false comprehension questions were presented at the end of 25% of the materials (probing both experimental manipulations and the filler

¹⁷ The temporal re-sequencing of events required when processing *before* is itself known to elicit LAN (Müntz et al, 1998), and therefore this type of sentence was not included in the experimental materials.

sentences); sentences without a comprehension question ended with a “Press for next” message that remained on the screen until the subject pressed either response button. Comprehension questions focused on the content of the preceding sentence, but crucially not any co-referential relationships contained within the sentence (e.g. “*Was Will deployed for a full year?*”). Comprehension questions appeared 2000 msec. after the end of the sentence-final word and remained on the screen until the participant made a response with one of two hand-held buttons. Response hands were balanced across participants to control for handedness, and the correct response was balanced across comprehension questions. The next stimulus sentence began 3000 msec. after the subject’s response. In order to familiarize participants with the task, they completed a ten-sentence practice session before beginning the experiment.

4.4 Electrophysiological recording

The electroencephalogram (EEG) was recorded from 26 tin electrodes mounted geodesically in a commercially available Electro-Cap. These sites included midline prefrontal (MiPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right medial frontal (LMFr and RMFr), left and right medial lateral frontal (LDFr and RDFr), left and right medial central (LMCe and RMCe), midline central (MiCe), left and right medial lateral central (LDCe and RDCe), left and right lateral temporal (LLTe and RLTe), left and right medial lateral parietal (LDPa and RDPa), midline parietal (MiPa), left and right lateral occipital (LLOc and RLOc), left and right medial occipital (LMOc and RMOc), and midline occipital (MiOc). Each electrode was referenced online to the reference electrode at the left mastoid and later re-referenced offline to the

average of the two mastoids. To monitor blinks and eye movements, electrodes were placed on the outer canthi and under each eye. Impedances were kept below $5K\Omega$ during recording. The EEG was amplified using James Long amplifiers with an online Band-pass filter (.01 to 100 Hz), and digitized at a sampling rate of 250 Hz.

4.5 Data analysis

The main analyses, as described below, focused on the critical main clause subject position (i.e. *he* in conditions A and B, and *Will* in conditions C and D in Table 1.1). Mean amplitude measurements were taken over 1000 msec. epochs (including a 100 msec. pre-stimulus baseline) with primary focus on the LAN time window between 300-500 msec. post-stimulus onset. Analyses between 500 and 800 msec. were also conducted in all comparisons, but were of special interest in the comparison of the two **Pronoun** conditions, due to the possibility that the antecedent-less pronoun might elicit late positivity, as in the studies discussed in section 1.6. When LAN effects in response to the critical word position appeared to extend beyond 800 msec., across several word positions, mean amplitude measurements of 2000 msec. epochs (including a 500 msec. pre-stimulus baseline) were taken. These multi-word averages enabled us to measure effects of sustained negativity across the course of the sentence. However, multi-word averages cannot distinguish between initial processing costs elicited by the first (critical) word in the average and additional processing costs contributed by subsequent words. For this reason, we followed King & Kutas (1995) and Phillips et al. (2005) in computing mean amplitude measurements relative to a 100 msec. baseline re-established *at each word position in the region spanned by the sustained effect*. If a sustained effect is due solely to an initial processing cost triggered by the first word in a

multi-word average that remains constant over time, there should be no observable differences in the responses to the individual words spanned by the sustained effect following the re-baselining procedure. On the other hand, if any subsequent word in the multi-word average contributes to the sustained effect, a significant difference should be evident in that particular individual word average. Trials contaminated by excessive muscle activity, amplifier blocking, or eye movements were discarded before averaging. This led to the rejection of 8.3% of trials in the single-word averages and 33.8% of trials in the multi-word averages¹⁸. The averaged data were algebraically re-referenced off line to the average of the activity recorded at the two mastoid sites. ERP waves were smoothed offline using a low-pass filter with a 7 Hz cutoff for visualization purposes only. Because of the design of the materials, a full factorial ANOVA was deemed inappropriate¹⁹. Rather, we first conducted one-tailed t-tests for predicted effects²⁰. For unpredicted effects, we used one-way ANOVAs to make the critical comparisons between (1) the two **Pronoun** conditions and between (2) the two **Name** conditions. Each analysis had the repeated measures of experimental condition (two levels) and electrode (26 levels). This will be referred to as the full analysis. In addition (independent of whether we first ran t-tests or omnibus ANOVAs), a subsequent distributional analysis was conducted for each of the three main comparisons, including

¹⁸ This high rejection rate was a consequence of averaging over the last four words (a period of two seconds) of sentences that were 15 words in length.

¹⁹ This type of analysis would force us to compare pronouns to proper names. This comparison would elicit differences driven solely by lexical factors of word class (Müller & Kutas, 1996) that were not the focus of this experiment.

²⁰ This type of analysis was conducted on effects that were predicted in terms of the sentence position at which they were elicited, the polarity of the effect and its distribution at the scalp (see section 1.6 for explicit predictions).

condition (two levels), hemisphere (left vs. right), anteriority (prefrontal vs. frontal vs. temporal vs. occipital), and laterality (lateral vs. medial). Electrodes included in the distributional analysis were left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right lateral temporal (LLTe and RLTe), left and right medial lateral parietal (LDPa and RDPa), left and right lateral occipital (LLOc and RLOc), and left and right medial occipital (LMOc and RMOc). This distributional analysis was conducted in order to investigate further those effects that appeared to be spatially localized. We felt that this type of analysis was warranted because one of the stated goals of the experiment was to examine the ways in which the brain's responses to referential and syntactic dependencies patterned together or differently, and one dimension along which these effects may differ is in their topographies over the scalp. Furthermore, when it was necessary to corroborate local effects suggested by the distributional analysis, ANOVAs were performed at individual levels of these factors. The Huynh-Feldt (1976) correction for lack of sphericity was applied, and corrected *p*-values are reported with the original degrees of freedom.

Chapter 1, in full, is a manuscript submitted to *Brain Research*, and is currently under review. Barkley, C., Kluender, R., & Kutas, M (*submitted*). The dissertation author was the primary investigator and preparer of this manuscript.

5. Works cited

Anderson, J., & Holcomb, P (2005). An Electrophysiological Investigation of the Effects of Coreference on Word Repetition and Synonymy. *Brain and Language*, **43**, 1-17.

Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. London: Routledge.

- Ben-Schachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The Neural Reality of Syntactic Transformations: Evidence from Functional Magnetic Resonance Imaging. *Psychological Science*, **14**(5), 433-440.
- Ben-Schachar, M., Palti, D., & Grozinsky, Y. (2004). Neural Correlates of Syntactic Movement: Converging Evidence from two fMRI experiments. *NeuroImage*, **21**, 1320-1336.
- Baddeley, A. (1992). Working Memory. *Science*, **255**(5044), 556-569.
- Caplan, D., & Waters, G. (1999) Verbal Working Memory and Sentence Comprehension. *Behavioral and Brain Sciences* 22, 77-126
- Caplan, D., Alpert, N., & Waters, G. (1998). Effect of Syntactic Structure and Propositional Number on Patterns of Regional Cerebral Blood Flow. *Journal of Cognitive Neuroscience*, **10**, 541-552.
- Caplan, D., Alpert, N., & Waters, G. (1999). PET Studies of Syntactic Processing with Auditory Presentation. *NeuroImage*, 43-51.
- Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of Broca's Area by Syntactic Processing under Conditions of Concurrent Articulation. *Human Brain Mapping*, **9**, 65-71.
- Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., & Dale, A. (2001). Vascular Responses to Syntactic Processing: Event-Related fMRI study of Relative Clauses. *Human Brain Mapping*, **15**, 26-38.
- Chomsky, N. (1980). On Binding. *Linguistic Inquiry*, **11**(1), 1-46.
- Chomsky, N. (1981). Lectures on Binding and Government. Foris Publications, Dordrecht.
- Chung, S. (1994). Wh-agreement and referentiality in Chamorro. *Linguistic Inquiry*, **25**, 1-44.
- Cinque, G. (1990). Types of A'-dependencies. Cambridge, MA: MIT Press
- Cooke, A. Zurif, E., DeVita, C., Alsop, D., Koenig, P., Detre, J., Gee, J., Pinango, N., Balogh, J. &
- Grossman, M. (2001). Neural Basis for Sentence Comprehension: Grammatical and Short-term Memory Components. *Human Brain Mapping*, **15**, 80-94.
- Coulson, S., King, J., & Kutas, M. (1998). Expect the Unexpected: Event-Related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, **13**(1), 21-58.
- Donchin, E. (1981). Surprise!.....Surprise? *Psychophysiology*, **18**(5), 493-513.
- Frazier, L (1979). On Comprehending Sentences: Syntactic Parsing Strategies. *University of Connecticut Dissertation*.
- Gundel, H., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring expressions. *Language*, **69**(2), 274-307.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The Syntactic Positive Shift (SPS) as an ERP Measure of Syntactic Processing. *Language and Cognitive Processes*, **8**(4), 489-483.

- Hagiwara, H., Soshi, T., Isihara, M., & Imanaka, K. (2007). A Topographical Study on the Event-Related Potential Correlates of Scrambled Word Order in Japanese Complex Sentences. *Journal of Cognitive Neuroscience*, **19**(2), 175-193.
- Harris, T., Wexler, K., & Holcomb, P. (2000). An ERP Investigation of Binding and Coreference. *Brain and Language*, **75**, 313-346.
- Heine, A., Tamm, S., Hoffman, M., Hutzler, F., and Jacobs, A. (2006). Does the Frequency of the Antecedent Noun Affect the Resolution of Pronominal Anaphors? An ERP Study. *Neuroscience Letters*, **400**, 7-12.
- Just, M. A., & Carpenter, P. A. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review*, **98**, 122-149.
- Just, M., Carpenter, P., & Keller, T., Eddy, W., & Thulborn, K. (1996). Brain Activation Modulated by Sentence Comprehension. *Science*, **274**, 114-116.
- Kaan, E., Harris, T., Gibson, E., & Holcomb, P. (2000). The P600 as an Index of Syntactic Integration Difficulty. *Language and Cognitive Processes*, **15**(2), 159-201.
- Kaan, E., Dallas, A.C., Barkley, C. M. (2007). Dealing with new discourse referents: An ERP study. *Brain Research*, **1146**, 199-209.
- Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The Effect of Syntactic Constraints on The Processing of Backwards Anaphora. *Journal of Memory and Language*, **56**, 384-409.
- King, J., & Kutas, M. (1995). Who Did What and When? Using Word- and Clause-Level ERPS to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, **7**(3), 376-395.
- Kluender, R., & Kutas, M. (1993a). Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. *Journal of Cognitive Neuroscience*, **5**(2), 196-214.
- Kluender, R., & Kutas, M. (1993b). Subjacency as a Processing Phenomenon. *Language and Cognitive Processes*, **8**(4), 573-633.
- Kok, A. (2001). On the Utility of P300 Amplitude as a Measure of Processing Capacity, *Psychophysiology*, **38**, 557-577.
- Kutas, M., & Hillyard, S. (1983). Event-related Brain Potentials to Grammatical Errors and Semantic Anomalies. *Memory and Cognition*, **11**(5), 539-550.
- Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics Electrified II (1994-2005). In M.A. Gernsbacher & M. Traxler (Eds.), *Handbook of Psycholinguistics*, 2nd edition. New York: Elsevier Press.
- Kwon, N., Kluender, R., Kutas, M., & Polinsky, M. (2013). Subject/object Processing Asymmetries in Korean Relative Clauses: Evidence from ERP data. *Language*, **89**(3), 537-585.
- Levin, B. (1993). English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press, London.
- Lewis, R., Vasishth, S., & Van Dyke, J. (2006). Computational Principles of Working Memory in Sentence Comprehension. *Trends in Cognitive Sciences*, **10**(10), 447-454.

- Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing Interpretable Discourse: The Establishment and Maintenance of Reference. In R. J. Jarvella & W. Klein (Eds.), *Speech, Place, and Action*. Chichester: Wiley.
- Martin, A.E., Nieuwland, M.S., & Carreiras, M. (2012). Event-related brain potentials index cue-based retrieval interference during sentence comprehension. *NeuroImage*, **59**, 1859-1869.
- Martin, A.E., Nieuwland, M.S., & Carreiras, M. (2014). Agreement attraction during comprehension of grammatical sentences: ERP evidence from ellipsis. *Brain and Language*, **135**, 42-51.
- Martin-Loeches, M., Munoz, F., & Casado, P. (2005). Are the Anterior Negativities to Grammatical Violations Indexing Working Memory? *Psychophysiology*, **42**(5), 508-519.
- Matzke, M., Mai, H., Nager, W., Russeler, J., & Münte, T. (2002). The Costs of Freedom: An ERP Study of Non-canonical Sentences. *Clinical Neurophysiology*, **113**, 844-852.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory Structures that Subserve Sentence Comprehension. *Journal of Memory and Language*, **48**(1), 67-91.
- Müller, H., & Kutas, M. (1996). What's in a name? Electrophysiological Differences between Spoken Nouns, Proper Names and one's Own Name. *Neuroreport*, **8**(1), 221-225.
- Müller, H., King, J., & Kutas, M. (1997). Event-Related Potentials Elicited by Spoken Relative Clauses. *Cognitive Brain Research*, **5**, 193-203.
- Münte, T., Schlitz, S., & Kutas, M. (1998). When Temporal Terms Belie Conceptual Order. *Nature*, **395**, 71-74.
- Neville, H., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991). Syntactically Based Sentence Processing Classes: Evidence from Event Related Brain Potentials. *Journal of Cognitive Neuroscience*, **3**(2), 151-165.
- Nieuwland, M., & van Berkum, J. (2006). Individual Differences and Contextual Bias in Pronoun Resolution. *Brain Research*, **1118**, 155-167.
- Nieuwland, M., Petersson, K.M., & Van Berkum, J. (2007a). On Sense and Reference: Examining the Functional Neuroanatomy of Referential Processing. *NeuroImage*, **37**, 993-1004.
- Nieuwland, M., Otten, M., & Van Berkum, J. (2007b). Who Are You Talking About? Tracking Discourse Level Referential Processing with Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, **19**(2), 228-236.
- Nieuwland (2014). "Who's he?" Event-related Brain Potentials and Unbound Pronouns. *Journal of Memory and Language*, **(76)**, 1-28.
- Osterhout, L., & Mobley, L. (1995). Event-related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*, **34**(6), 739-773.
- Pesetsky, D. (1990). Wh-in-situ: Movement and Selective Binding. In Reuland and Martin (eds.), *The representation of (in-)definiteness*. Cambridge, MA: MIT Press.
- Phillips, C., Kazanina, N., & Abada, S. (2005). ERP Effects of the Processing of Syntactic Long-distance Dependencies. *Cognitive Brain Research*, **22**, 407-428.

- Reinhardt, T., & Reuland, E. (1993). Reflexivity. *Journal of Linguistic Inquiry*, **24**, 657-720.
- Rose, S. (2011). Long Distance Assimilation of Consonants. *The Blackwell Companion to Phonology*, Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (eds.). Wiley-Blackwell.
- Rose, S., & Walker, R. (2011). Harmony Systems. In J. Goldsmith, J. Riggall & A. Yu (eds.) *Handbook of Phonological Theory* (2nd ed.). Blackwell.
- Rösler, F., Pechmann, T., Streb, J., Roder, B., & Henninghausen (1998). Parsing of Sentences in a Language with Varying Word Order: Word-by-word Variations of Processing Demands are Revealed by Event-Related Potentials. *Journal of Memory and Language*, **38(2)**, 150-176.
- Ruchkin, J., Johnson, R., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: an event-related potential study. *Cognitive Brain Research*, **1(1)**, 53-66.
- Shao, J., & Neville, H. (1998). Analyzing Semantic Processing Using Event-Related Brain Potentials. In *The Newsletter of the Center for Research in Language*.
- Silverstein, M. (1976). Hierarchy of Features and Ergativity. In R. M. W. Dixon (Ed.), *Grammatical Categories in Australian Languages*. New Jersey: Humanities Press
- Streb, J., Henninghausen, E., & Rösler, F. (2004). Different Anaphoric Expressions are Investigated by Event-Related Brain Potentials. *Journal of Psycholinguistic Research*, **3**, 175-201.
- Stromswold, K., Caplan, D., Alpert, N., & Rausch, S. (1996). Localization of Syntactic Comprehension by Positron Emission Tomography. *Brain and Language*, **52**, 452-473.
- Ueno, M. & Garnsey, S. (2007) An ERP Study of the Processing of Subject and Object Relative Clauses in Japanese. *Language and Cognitive Processes*, **23(5)**, 646-688.
- Ueno, M., & Kluender, R. (2003). Event-related Brain Indices of Japanese Scrambling. *Brain and Language*, **86**, 243-271.
- Ueno, M., & Kluender, R. (2009) On the Processing of Japanese Wh-questions. *Brain Research*, 1-29.
- Van Berkum, J., Brown, C., & Hagoort, P. (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials. *Journal of Memory and Language*, **41**, 147-182.
- Van Berkum, J., Brown, C., Hagoort, P., & Zwisterlood, P. (2003). Event-Related Brain Potentials Reflect Discourse-Referential Ambiguity in Spoken Language Comprehension. *Psychophysiology*, **40**, 235-248.
- Berkum, J. (2004). Sentence Comprehension in a Wider Discourse: Can we Use ERPs to Keep Track of Things? In M. Carreiras, Jr. & C. Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York: Psychology Press.
- Van Berkum, J., Koorneef, A., Otten, M., & Nieuwland, M. (2007). Establishing Reference in Language Comprehension: An Electrophysiological Perspective. *Brain Research*, **1146**, 158-171.
- Vonk, W., Hustinx, L., & Simons, W. (1992). The Use of Referential Expressions in Structuring Discourse. *Journal of Verbal Learning and Verbal Behavior*, **18**, 545-567.

Weckerly, J., & Kutas, M. (1999). An Electrophysiological Analysis of Animacy Effects in the Processing of Object Relatives. *Psychophysiology*, **36**, 559-570.

CHAPTER 2: The Role of Selective Attention in Implicit Language Learning: An Event-related Potential (ERP) Study

Abstract:

A substantial body of research investigating implicit language learning at various levels of linguistic analysis has proven it to be a powerful and flexible learning mechanism that enables language learners to acquire striking amounts of linguistic information quite rapidly merely during passive exposure. However, the precise role of selective attention in implicit language learning has yet to be systematically characterized: previous studies (Saffran et al. 1997, Toro et al. 2005) have reached opposite conclusions using word segmentation tasks in dual-task settings, but no study has yet investigated the role of selective attention in implicit learning tasks involving higher levels of linguistic analysis. To address this issue, we constructed an artificial language and used a traditional training-testing paradigm to determine whether unattended morpho-syntactic information can be implicitly learned. During training, participants were instructed to attend to the verbal agreement system of the language with the explicit goal of learning it; unbeknownst to them, the verb phrase also contained information about the transitivity of the matrix verb and the auxiliary that it selected for. During testing, while their EEG was recorded, participants judged the grammaticality of correct sentences and sentences containing violations of both the attended and unattended systems. Performance during both training and testing demonstrated rapid learning of the regularities governing the attended agreement system, while participants' ability to detect unattended transitivity and auxiliary selection violations during testing was below chance. Nevertheless, all three violation types elicited brain responses that differed significantly from the responses to correct

sentences. Correlational analyses showed further that performance in mastering the agreement system during training correlated negatively with the ability to detect transitivity and agreement violations during testing. The ERP results go beyond previous demonstrations of brain/behavior dissociations in explicit adult language learning (McLaughlin et al. 2004, Tokowicz & McWhinney 2005) by demonstrating a brain/behavior dissociation in the acquisition of morpho-syntactic features outside the focus of attention within a period of a few hours. We conclude by discussing the implications of these results for adult second language acquisition.

1. Introduction

Acquiring a language is an immensely complex endeavor. The language learner must construct an intricate cognitive system characterized by complexity at all levels, yet the language input into this learning process is continuous, perceptually degraded, and often under-representative of the linguistic systems to be learned. In spite of this, language is acquired effortlessly, instinctively, and with almost breathtaking speed. By 4 days of age, infants can discriminate between their native language and others on the basis of prosody (Mehler et al., 1998), at 4 months of age can exploit statistical information to segment the continuous speech stream (Saffran et al., 1996) and by between 9 and 12 months of age have acquired the rudimentary phonological (Jusczyk et al., 1993; Jusczyk et al., 1999) and phonotactic structure (Jusczyk et al., 1993) of their native language. Additionally, during their first year of life, their linguistic environment has shaped how they perceive speech (Werker & Tees, 1984), and they have begun the lifelong process of building a mental lexicon (Tincoff & Juczyk, 1999).

Researchers in the field of language acquisition have long sought to determine the cognitive mechanisms that underlie these impressive feats.

So-called *experience-independent* theories initially dominated the field of language acquisition. Proponents argued, often building on the conceptual foundation of Universal Grammar (Chomsky, 1965) that language acquisition proceeds so effortlessly because children approach the task with innate, hard-wired linguistic knowledge. This linguistic endowment is claimed to guide the early learning process and to explain how learners overcome the “poverty of the stimulus” (Chomsky, 1980) and acquire language with such speed. In opposition to this are *experience-dependent* accounts, which posit that humans possess powerful domain-general learning mechanisms that are recruited during language acquisition, eliminating the need for guidance by innate linguistic knowledge (see for example Eimas et al., 1971 and Saffran et al., 1996 among many others). Perhaps the most often discussed of these learning mechanisms, and the focus of the present work, is implicit learning, and specifically the role that attention plays in this learning process.

Implicit learning,¹ a term first appearing in Reber (1967), is learning that is automatic, incidental, and spontaneous. This type of learning occurs absent any conscious intent to learn,² does not require the formation or testing of hypotheses, and

¹ Implicit learning is often also referred to as statistical, or distributional, learning. Some researchers argue that these constitute distinct learning mechanisms that are best addressed with distinct experimental approaches; see Perruchet & Pacton (2006) for a recent unifying perspective. For the sake of clarity, here we adopt the term implicit learning to the exclusion of the others, with the non-ideological intent that this term be taken in the general sense, rather than in opposition to putative statistical or distributional learning.

² Implicit learning has even been observed in rats (Toro & Trobalon, 2005) and non-human primates (Hauser, Newport, & Aslin, 2001) using the types of experimental paradigms discussed here; these species could presumably not be imbued with any sort of “conscious” intent to learn.

both the learning mechanism itself and its (behavioral) output often remain below the level of conscious awareness. Passive exposure to a system governed by systematic regularities is sufficient for implicit learning to occur. This learning mechanism is perfectly suited for language acquisition, as it can take advantage of the rich structure inherent in linguistic input by extracting vast amounts of information without the learner having to consciously control the learning process (see for example Saffran et al., 1996).

Implicit learning experiments generally consist of familiarization and testing phases, and it is this general approach that will be adopted here. During familiarization, a highly structured stimulus set is presented to participants in the absence of any explicit instructions to learn (i.e. participants often simply listen to/view the stimuli passively). After familiarization, learning of the regularities present in the input is assessed, usually by means of an offline two-alternative forced-choice task (2AFC) that requires participants to select between stimuli consistent with the regularities of the input and foil stimuli. Employing variations of this basic paradigm, a multitude of implicit learning studies have investigated the learning of various types of (simple and complex) statistical regularities that govern the relationships between items in the learning set, as discussed in sections 1.1 – 1.5.

While the implicit learning literature is vast and a comprehensive review is beyond the scope of this article, there is a comparatively small literature on the role that attention plays in implicit learning, and implicit language-learning in particular. Below we briefly discuss studies examining implicit learning of linguistic information, before turning to a discussion of studies that employ manipulations of attention in tasks that

assess learning of both non-linguistic and linguistic stimuli. Lastly, we discuss ERP studies of explicit language-learning, as the goal of the current study was to extend these lines of research by recording ERPs while employing an implicit language-learning paradigm with a novel attentional manipulation, an experimental design that enabled us to investigate the interplay between attention and implicit language learning using a temporally fine-grained experimental technique.

1.1 Implicit learning of language

The claim that implicit learning plays a role in natural language acquisition often cites Saffran et al., (1996) as its empirical foundation. This seminal study set off an explosion of subsequent research, a subset of which is reviewed below. Saffran et al. (1996) exposed eight-month-old infants³ to a two-minute speech stream consisting of random repetitions of four three-syllable nonsense words. Crucially, transitional probabilities between syllables were high within “words,” but low across word boundaries. If participants are able to extract this information and use it to segment words from the “speech” stream, then different behavioral responses to “words” and “non-words” would be expected after familiarization. Using a preferential looking paradigm (Hirsh-Pasek & Golinkoff, 1996), this is exactly what the authors reported: in two experiments, participants spent significantly longer looking in the direction of “non-words” compared to “words” from the speech stream (the “novelty preference”), regardless of whether the “non-words” never appeared during familiarization, or did but only when crossing “word” boundaries, i.e. with lower transitional probabilities.

³ Participants in the studies reviewed here ranged from infancy to adulthood. Implicit learning in particular, and processes of incidental frequency estimation in general, are argued to be age-invariant, with the performance of young children and adults basically equivalent (see Hasher & Chromiak, 1977; Ellis, Palmer, & Reeves, 1988; and Reber, 1993).

Building on this methodology, a series of studies have explored whether transitional probability information can be used for processes of word segmentation, or the acquisition of information at the level of phrase-structure and the lexicalized properties of verbs (Saffran, 2001; Saffran, 2002; Saffran & Wilson, 2003; Thompson & Newport, 2007; Wonnacott et al., 2008).

The experiments reported in Saffran (2001a), Saffran (2002), Saffran & Wilson (2003), and Thompson & Newport (2007) all build phrase structure into artificial languages in essentially the same fashion. Nonsense words are assigned to grammatical classes, and the syntax of the language is characterized by the statistical distributions of these categories. Saffran (2001a, 2002) showed that both infant and adult learners are indeed sensitive to this information, as again participants performed above chance on a post-familiarization 2AFC task similar to that used in the word-segmentation task described above. Thompson & Newport (2007) replicated these findings employing a range of artificial languages with varying degrees of complexity, and Saffran & Wilson (2003) linked the results of word- and phrase-level studies by demonstrating that after words are segmented from the speech stream, this lexical information can subsequently feed processes of phrasal grouping.

Though the studies above appear to demonstrate the power of implicit learning, objections have been raised. Researchers have argued that implicit learning may simply be an artifact of artificial laboratory settings, unnatural linguistic input, and/or the use of offline dependent measures. Saffran (2001b) tested the hypothesis that participants are simply extracting chunks of highly coherent “sound units” from a “speech” stream, and then applying this information to yield high accuracy on a 2AFC task, without affording

these sequences real linguistic status (in which case one could argue that implicit learning does not underpin the process of language acquisition/learning). Saffran tested this hypothesis using the familiar 2AFC paradigm, but placed the two choices in either English or nonsense carrier phrases. She found that participants only displayed the novelty preference when words were placed in English carrier phrases, leading her to argue that the output of implicit learning does indeed have actual linguistic status for the learner.

The experiments in Wonnacott et al., (2008) merit special attention here, as the research described therein addressed the issue of whether implicit learning yields representations that guide online language comprehension in more naturalistic settings. Data were collected using online measures, nouns had referents, and visual scenes accompanied sentence presentation. The authors used an artificial language to determine whether the acquisition of verbal subcategorization information can be accounted for with implicit learning.⁴ After familiarization, performance on three tasks revealed that the participants had acquired the sub-categorization system of the language: they avoided over-generalizing during a production task, made accurate grammaticality judgments, and, most critically, eye-tracking results indicated that the participants were using their acquired knowledge of verbal subcategorization to guide comprehension in real time.

In sum, these studies show that the implicit learning mechanism is powerful, flexible, and can operate over linguistic input to extract structured information at

⁴ An artificial language allows the authors to instantiate the argument structures of individual verbs in a purely statistical fashion, avoiding confounds (e.g., semantics) inherent in trying to investigate this issue with natural language materials. We adopt the same approach here for this very reason.

multiple linguistic levels. As stated above, the primary goal of the present study is to extend this research by investigating the role that attention plays in this process.

1.2 Implicit learning and attention

There are limits to the amount of information that humans can process at any given time, both perceptually and cognitively. Due to these limitations, a neurocognitive mechanism is required that can allocate processing resources and select the aspects of the environment and cognitive tasks that have priority for the organism. This mechanism is attention. Attention research has a long history, but there is an important conceptual bifurcation (of an admittedly multi-faceted construct) to be made that will facilitate our exposition here, helping provide an account of why the role of attention in implicit learning has yet to be consistently characterized. Since at least Johnston & Dark (1986), the field has been dominated by two main views of attention: attention as resource, and attention as selection mechanism.

The view of attention as a resource follows from the assumption that (non-automatic) processes that require attention must draw these resources from a limited central resource pool. It then follows that performing multiple tasks simultaneously will increase the draw on this pool, until concurrent processing is sufficient to deplete resources entirely, resulting in decrements in performance. This conceptualization of attention motivates the design of studies that compare performance in single- and dual-task settings, with the goal of probing the attentional resources required by various processes.

The view of attention as a selection mechanism operates under the assumption that not all input to the sensory periphery can be processed efficiently, and that attention

functions to select the input that will be processed fully. In other words, properties of the input that are selected by this mechanism are processed preferentially. This view of attention (which we will subsequently refer to as *selective attention*) has led to a number of alternative selective attention manipulations, in which task parameters dictate that only a subset of the stimulus display need be attended for successful task performance. It is then possible to assess the effects of this manipulation on the learning of the “to-be-attended” and “to-be-ignored” aspects of the stimulus. It is this type of attentional manipulation that we adopt here, a manipulation that has no precedent in the language-learning literature.

Though this distinction is important within the attention literature and has motivated a substantial amount of research testing the predictions of each account, providing support for one of these views of attention is not the focus of the current work as we employ a manipulation of selective attention in lieu of an experimental design comparing performance under “single-task settings” and “dual-task settings.” We simply mention these competing views here in order to operationalize the term “selective attention” and because they have motivated the use of very different experimental paradigms, the conflicting results of which have greatly contributed to a lack of clarity with regards to the role of attention in implicit learning.

1.3 The role of attention in the implicit learning of non-linguistic information

In an oft-cited study, Nissen & Bullemer (1987) pioneered the idea that implicit learning may be dependent on attentional resources (i.e. the “attention as resource” view

referred to in the previous section). Using a serial reaction time task (SRT),⁵ the authors compared training performance under single-task conditions to performance under dual-task conditions, in which subjects were required to perform a secondary task. In the single-task conditions, the authors reported evidence of learning: subjects showed a decrease in response times when presented with structured input but virtually no decrease in RT when the input was random. No learning was observed in the dual-task conditions, however, suggesting that dual-task conditions not only interfered with implicit learning but in fact blocked it entirely. In other words, these dual task conditions appeared to sufficiently deplete the pool of attentional resources to the point that the implicit learning mechanism was unable to function efficiently.

The experiments in Jimenez & Mendez (1999) represented an important advance in the nature of the experimental manipulation designed to tax attention, which is of particular pertinence to the current study. They again employed the SRT, but built two sets of regularities into the stimuli: both the location *and* shape of the current stimulus provided predictive information about the subsequent stimulus. One group of participants simply performed the SRT task (i.e. shape information was task-irrelevant), while a second group was instructed to track the occurrence of certain shapes (in addition to stimulus locations) throughout the experiment. If implicit learning is always blocked by an increase in attentional load, then only the first group of participants should show evidence of learning. However, if implicit learning is dependent on participants selectively attending the aspect of the stimulus encoding the to-be-learned

⁵ See Nissen & Bullemer (1987) for a detailed description of the structure of this task.

pattern, then only the second group should show evidence of learning the mapping between shape and location. This is exactly the pattern that was observed, suggesting that it is not increases in attentional load *per se* that block implicit learning. Rather, it seems that, in an input set containing stimuli that vary systematically along multiple dimensions, the regularities attended to are the ones that are learned, suggesting a central role for selective attention in implicit learning.⁶ The present study utilizes a similar manipulation in order to determine whether the same holds true for linguistic input while also assessing the effects of selective attention manipulations using sensitive online measures.

1.4 The role of attention in implicit language learning

The experiments in Saffran et al. (1997) were very similar to those in Saffran et al. (1996) in that they centered around processes of word segmentation, with the additional inclusion of an attentional manipulation. To deplete attentional resources, participants created illustrations on a computer during exposure to the type of speech stream described above. Across two experiments, participants performed above chance in this dual-task setting, with the authors interpreting these results as indicating that implicit learning is incidental and not dependent on attention, and can take place even in cases when participants are otherwise engaged.

Toro et al., (2005) further investigated this issue with a more systematic manipulation of attention. The authors argued that the data in Saffran et al. (1997) were

⁶ See Jiang & Chun (2001), Baker, Olson, & Behrmann (2004), and Hoffman & Seibald (2005) for additional studies investigating the role of selective attention in implicit learning using visual search tasks. By manipulating the task relevance of the pattern to be learned, the authors were again able to influence whether implicit learning took place.

inconclusive because (i) there was no guarantee that participants entirely ignored the speech stream in the dual-task setting, (ii) there is evidence that task-irrelevant information may undergo some processing regardless of instruction if the main task is not sufficiently demanding (as the coloring task used in Saffran et al. (1997) presumably was not), and (iii) performance did appear to be somewhat degraded in comparison to previous studies that did not manipulate attention. After changing the nature of the secondary task, Toro et al. (2005) observed differences in performance under single- and dual-task settings, arguing that manipulations of attention do have at least some effect on implicit language-learning, contra the claims in Saffran et al. (1997). However due to the dual-task design of these two experiments, the results shed no light on the role that *selective* attention plays in implicit learning of linguistic input, an issue the current is designed to investigate.

Collectively, these results paint a complex picture of the role of attention in implicit learning. Manipulations of attentional load using dual-task paradigms have been observed to both influence (Nissen & Bullemer, 1987; Toro et al., 2005) and not influence (Saffran et al., 1997) implicit learning in offline tasks, and therefore no straightforward conclusions can be drawn as to the role of attentional resources in this type of learning. In contrast, manipulations of selective attention seem to disrupt implicit learning consistently, regardless of the task employed (Jimenez & Mendez, 1999; Jiang & Chun, 2001; Baker, Olson, & Behrmann 2004; Hoffman & Sebold, 2005). While a comprehensive characterization of the influence of attention on implicit learning mechanisms is thus still lacking, it is clear that this relationship has been much more comprehensively explored in the non-language literature than it has in the

corresponding language literature, which lacks studies containing systematic manipulations of selective attention.

As such, the primary goal of the current study is to manipulate selective attention systematically during a language-learning task in order to investigate the role of attention in implicit language learning, specifically at higher levels of linguistic analysis. In doing so, we aim to determine whether manipulations of selective attention affect implicit *language* learning in the same fashion as in the non-linguistic studies described above (Jimenez & Mendez, 1999; Jiang & Chun, 2001; Baker, Olson, & Behrmann 2004; Hoffman & Seibald, 2005), and in doing so fill a gap in the existing literature on attentional influences on implicit language learning. Additionally, our novel use of the ERP technique, (in addition to behavioral measures), a multi-dimensional measure, has the potential to reveal evidence of implicit learning of unattended linguistic information that has thus far remained elusive.

1.5 Event-related potential (ERP) studies of (artificial) language learning

In the current study, as in the studies described below, we utilized ERPs, an experimental technique that enables us to examine the time course of processing with a high degree of temporal precision, thus allowing us to look for evidence of implicit learning that offline behavioral measures might be insufficiently sensitive to detect. Friederici et al. (2002) conducted an ERP experiment examining implicit learning of an artificial grammar, and provided additional support for the position that implicit learning underlies naturalistic language learning and guides online comprehension. The authors constructed a grammar using a small set of phrase structure rules and trained adult participants within the setting of a game, teaching them how to articulate all of the

game's moves with the artificial language, ensuring that the language was learned in an environment in which it served a communicative purpose. After grammatical training, ERP indices of grammatical processing difficulty in response to violations of the phrase structure rules of this language were both early (approximately 100 msec. after the onset of the ungrammaticality) and sustained. While the functional interpretation of these brain responses is largely tangential, it is worth pointing out that the observed ERP effects are traditionally interpreted as syntactic in nature, and are often elicited by natural language materials that engender a problem with parsing the input into a grammatical sequence.

McLaughlin et al. (2004) conducted a study examining lexical acquisition in native English speakers learning French as a second language (L2), using N400 responses and a lexical decision task as dependent measures. For native speakers, the N400 response is largest in amplitude to phonologically legal non-words, intermediate in amplitude when a word is not primed, and smallest in amplitude when a target word is preceded by an associated prime. Strikingly, at the first experimental session, after an average of only 14 hours of classroom instruction, the amplitude of the N400 response of L2 learners of French was already smaller to words taken directly from their textbook than to words from the same textbook in which one or two of the internal letters had been replaced – even though performance on the lexical decision task using these same words and pseudowords was at chance (as measured by d').⁷ This pattern of results suggests that learners are rapidly extracting lexical knowledge, but that this knowledge

⁷ By the second session, after an average of 63 hours of instruction, L2 learner N400 responses were also smaller to primed than to unprimed words taken from their textbook, essentially replicating the L1 pattern of N400 responses.

is not sufficiently robust to influence behavioral decisions. These results suggest that early evidence of explicit learning after brief exposure may be detectable only at the level of brain responses. This in part motivated our decision to use ERPs as a dependent measure and enabled us to address an additional hypothesis outlined below.

Tokowicz & MacWhinney (2005) used ERPs and grammaticality judgments to study the acquisition of grammatical features by native English speakers learning Spanish as an L2. Participants read grammatical and ungrammatical Spanish sentences and made end-of-sentence grammaticality judgments while their brain responses were recorded. Violation types included omission of required auxiliary verbs (similar across L1 and L2), sentences that violated patterns of determiner-number agreement (a grammatical system present in the L1 but with different constraints than in the L2), and sentences that violated determiner-gender agreement (a grammatical system present in the L2 but not in the L1). Violations of the auxiliary and gender-agreement systems elicited effects of late positivity between 500 and 900 msec. (P600 effects), even though grammaticality judgment accuracy was at chance for all violation types (as measured by one-tailed t-tests and d'), and especially for gender-agreement violations. This suggested that although at some level participants were sensitive to grammatical violations in the L2, this sensitivity was not sufficiently robust to influence their behavioral decisions.⁸

Building on the studies described above, the present study aims to introduce a

⁸ See Osterhout et al. (2006, 2008), McLaughlin et al. (2010) and Tanner et al. (2013) for related ERP studies of natural L2 learning showing the progression of N400 and P600 effects that emerge during the process of explicit learning. As our focus is on implicit learning of artificial languages, these studies will not be discussed in detail here.

novel⁹ implicit learning paradigm into an ERP study of (artificial) language learning, while also including a selective attention manipulation that has no precedent in the ERP language-learning literature. If we can in fact obtain evidence of implicit learning of unattended grammatical information despite *very* brief exposure (see section 2.3.1), we believe that we will have demonstrated two important points: first, that selective attention is not a pre-requisite for implicit language learning, and second, that the early outputs of learning (may) be detectable only in brain responses, perhaps shedding light on observations, in the form of behavioral responses, that unattended non-linguistic regularities can not be implicitly learned. If this pattern of results does in fact obtain, it would cast light on the contradictory findings discussed above and bridge the gap between ERP studies of explicit and implicit language learning and studies investigating the role of selective attention in implicit learning of non-linguistic information.

2. Methods

The present study was designed to investigate the role that selective attention plays in the implicit learning of grammatical systems. Using acceptability judgments and ERPs, we tested whether it is possible to learn elements of a grammatical system when attention is directed away from them. In order to test this hypothesis, we designed an experiment with training and testing phases in which the participant's task was to learn the agreement system of an artificial language created for specifically for the experiment. The explicit instructions provided to participants during the training phase

⁹ Though see Kutas & Baldwin (1993) for an ERP study comparing implicit versus explicit learning of a finite state grammar. Their results suggest that participants did in fact extract some regularities under implicit learning conditions. We do not focus on this study here as our focus is on implicit learning at higher levels of linguistic analysis, namely morpho-syntactic information.

of the experiment were designed to ensure that they directed their attention to this agreement system with the goal of learning it. Unbeknownst to participants, however, the training materials included two additional grammatical systems that encoded constraints on verbal subcategorization at the exact same sentence positions as the agreement system, namely in the verb phrase: verbs were either transitive or intransitive, and selected for lexically specified auxiliary verbs. Though these distinctions were marked at the same sentence positions as the agreement system, this grammatical information was entirely task-irrelevant and independent of the agreement system, and therefore successful task performance did not depend to any extent on learning the regularities governing these additional grammatical systems. This manipulation, along with the training-testing design of our experiment, made it possible for us to assess the role of selective attention in the acquisition of attended (agreement) and unattended grammatical systems (transitivity and auxiliary selection) after brief language exposure.

2.1 Materials

In order to address our research questions adequately, we created an artificial language¹⁰ from scratch that was comparable in complexity to naturally occurring languages, opting against an overly simplified or overly complex system that may have yielded floor or ceiling effects during training. With this goal in mind, we created an artificial language with SOV word order and a split ergative case marking/number

¹⁰ We would like to thank Maria Polinsky for advising us on the particular typological features of this language.

agreement system¹¹ (in which case was in ergative-absolutive alignment and number marking in nominative-accusative alignment). The language contained ten nouns, eight matrix verbs (four transitive, four intransitive) and four auxiliary verbs. Matrix verbs were prefixed with an agreement marker (e.g. **kig**-glim in condition A (intransitive construction) and **kib**-bund in condition E (transitive construction) in Table 1 below) that indicated case and number agreement with the subject. Matrix verbs also selected for a specific, obligatory and lexically specified auxiliary verb, two of which were selected only by transitive verbs and the other two only by intransitive verbs. Example sentences are provided in TABLE 2.1 below, and a complete description of the artificial language can be found in Chapter 2, Appendix 1.

2.1.1 Materials: Training session

During the session, participants were exposed only to grammatical sentences (conditions A and E). Sentences were presented word-by-word until the main verb was encountered, at which point participants were presented with a 2AFC task (see section 2.3). In this task, they selected the matrix verb that obeyed the constraints of the agreement system they were being trained on, and which therefore rendered the sentence grammatical. In other words, their task was to discriminate between agreement markers on matrix verbs that were grammatical (TABLE 2.1, conditions A [**kig**-glim] and E [**kib**-bund]) or ungrammatical (TABLE 2.1, conditions B [***kig**-glim] and F [***kab**-bund]) in the context of the entire sentence. Crucially, during training, participants were not exposed to any violations of the transitivity (TABLE 2.1, conditions C and G) or auxiliary selection systems (TABLE 2.1, conditions D and H).

¹¹ See Dixon (1994) for an extensive discussion of split ergativity.

TABLE 2.1 Example sentences of the artificial language used during training and testing sessions¹²**Intransitive verb selecting auxiliary 1, singular agreement with ABS subject****A. Grammatical**

DET	[SUBJ.NP]ABS/SING	AUX.1	SING[V-INTRANS]
Ba	tombat-il	til	kig-glim

B. *Agreement violation

DET	[SUBJ.NP]ABS/SING	AUX.1	SING[V-INTRANS]
Ba	tombat-il	til	*kig-glim

C. *Transitivity violation

DET	[SUBJ.NP]ABS/SING	AUX.1	SING[V-INTRANS]
Ba	tombat-il	til	kib-*bund

D. *Auxiliary selection violation

DET	[SUBJ.NP]ABS/SING	AUX.1	SING[V-INTRANS]
Ba	tombat-il	*dal	kig-glim

Transitive verb selecting auxiliary 2, plural agreement with ERG subject**E. Grammatical**

DET	[SUBJ.NP]ERG/SING	[OBJ.NP]ABS	AUX.1	SING[V-TRANS]
Ba	tombat-e	blergen-amt/il	iti	ke-bund

F. *Agreement violation

DET	[SUBJ.NP]ERG/SING	[OBJ.NP]ABS	AUX.1	SING[V-TRANS]
Ba	tombat-e	blergen-amt/il ¹³	iti	*ko-bund

G. *Transitivity violation

DET	[SUBJ.NP]ERG/SING	[OBJ.NP]ABS	AUX.1	SING[V-TRANS]
Ba	tombat-e	blergen-amt/il	iti	kig-*glim

H. *Auxiliary selection violation

DET	[SUBJ.NP]ERG/SING	[OBJ.NP]ABS	AUX.1	SING[V-INTRANS]
Ba	tombat-e	blergen-amt/il	*ada	ke-bund

¹² Note that during the training session, subjects saw only grammatical sentences up to the verb phrase, at which point they made a binary choice between a main verb that either was marked with a licit agreement marker or one that rendered the sentence ungrammatical.

¹³ Note that in transitive structures, the object noun phrase is not participating in an agreement relationship, and therefore either singular or plural absolutive markers lead to licit agreement.

2.1 Materials: Testing session

During the testing session, participants read sentences that fully conformed to the grammar as described above (TABLE 2.1, conditions A and E), sentences that violated the attended agreement system (TABLE 2.1, conditions B and F), and violations of both the unattended transitivity (TABLE 2.1, conditions C and G) and auxiliary selection systems (TABLE 2.1, conditions D and G). Focusing on the intransitive structures for ease of exposition, in the agreement violations (TABLE 2.1, condition B) the sentence is rendered ungrammatical by the appearance of the singular marker *keg-* on the main verb *glim* in place of the plural marker *kig-*, thereby creating a number-marking mismatch between subject and verb. The transitivity violation (TABLE 2.1, condition C) is caused by the appearance of the transitive verb *bund* in an intransitive structure. Lastly, the auxiliary selection violation (TABLE 2.1, condition D) is generated by the appearance of the auxiliary, *dal* which does appear with intransitive verbs, but not with this specific intransitive matrix verb. Crucially, the agreement relations in the transitivity and auxiliary violations were licensed by the artificial language, such that participants never read sentences containing more than one violation. As detailed above, this ensured that participants were exposed to all aspects of the artificial language's grammar while the task that they performed directed their attention only to the agreement system.

2.2 Participants

20 monolingual native speakers of English (9 females, 11 males) participated in the experiment for course credit or monetary compensation. All participants were between the ages of 18-24 (mean 20.7) and enrolled as students at UCSD. All

participants were right-handed with no neurological or psychiatric disorders and normal or corrected-to-normal vision.

2.3 Procedure

2.3.1 Behavioral training session

At the beginning of the training session, prior to the commencement of the explicit grammatical learning task, participants were trained on the labels for the ten nouns in the artificial language (each of which referred to an image of an alien creature) and the meanings of four temporal adverbials. We trained participants on the meaning of these adverbials because they occurred sentence finally in our materials, allowing us to avoid potential sentence wrap-up effects that may obscure brain responses to the critical main verb position. Participants were then tested to criterion on these nouns and adverbs and instructed that they would be tested again at the end of the session. This was done simply to ensure that participants were processing the language stimuli at some level of depth and not treating the materials as entirely non-linguistic. Part of our primary experimental manipulation depended on the subcategorization frames of individual verbs (namely the transitivity and auxiliary selection systems), but we did not train participants on verb meaning or verbal subcategorization frames for two reasons. First, the verb's argument structure in English could be a source of interference as participants (potentially) implicitly learned the subcategorization frames of the individual verbs in our artificial language. Second, by refraining from training participants on subcategorization, we ensured that they would not be alerted to the presence of regularities governing the distribution of main and auxiliary verbs in the artificial language that they were intentionally not trained on.

At the end of this brief lexical training session, participants were seated comfortably in a quiet room in front of a PC running PsyScope^(c) and equipped with a button box so that they could be trained on the artificial language's agreement system. Participants were given instructions about the artificial grammar-learning task and told that they were to focus solely on learning the language's agreement system. The task was explained to the participants using pertinent examples from English, namely examples of licit (e.g. *he walks*) and illicit (e.g. **he walk*) subject-verb agreement, in order to ensure that they understood the task. Participants then performed 384 trials of a self-paced two-2AFC task that lasted approximately 45-60 minutes. Sentences were visually presented above fixation, with each word presented for a duration of 300 msec. and with a stimulus onset asynchrony (SOA) of 500 msec. This mode of presentation continued until participants reached the verb phrase, at which position the experimental manipulation had been encoded.¹⁴ At this point, participants were presented with two verb phrases (including the auxiliary) that differed only in their agreement markers (for example, "*kig-glim* or *kog-glim*?" occurring after a sentence containing a subject such as *tombat-il* (see TABLE 2.1, conditions A and B), with the goal of selecting the verb phrase that was marked with the grammatical agreement marker. In the case of an incorrect response, participants received an "incorrect" message and were then presented with the entire sentence containing the grammatical verb phrase, which remained on screen without further comment until a response button was pressed. In the case of a correct response, participants simply received a "correct" message. Pressing a response button after the feedback message was displayed triggered the presentation of

¹⁴ We selected this mode of presentation in order to be consistent with the ERP testing session.

the next sentence. In order to prevent a simple guessing strategy and to provide motivation throughout training, participants were informed before the experiment that there would be a small prize for the subject receiving the highest score during the training phase. A running tally of individual participants' correct responses was then continually updated during the training session in the top right hand corner of the screen, adjacent to the highest score in the experiment thus far. This strategy proved highly effective in engaging participants.

Most importantly for our investigation into the role of selective attention in implicit learning, the unattended verbal transitivity and auxiliary selection systems were also encoded in the verb phrase: main verbs occurred in either transitive or intransitive structures and occurred only with lexically specified auxiliary verbs, which were also divided into transitive and intransitive classes. However, though the participants were passively exposed to this system while attending to the language's agreement markers, this grammatical information was entirely task-irrelevant and therefore remained unattended throughout the course of the training session.¹⁵ This manipulation enabled us to compare the acquisition of grammatical systems that were either attended (agreement) or unattended (transitivity and auxiliary selection) during language exposure.

2.3.2 Testing session

During this portion of the experiment, the participants' task was to judge the grammaticality of legal and illegal sequences that violated grammatical systems both

¹⁵ Post-experiment debriefing questionnaires confirmed that participants were not consciously aware of the existence of the regularities governing verbal transitivity and auxiliary selection during training.

attended (agreement) and unattended (auxiliary selection and transitivity) during training (see TABLE 2.1). Participants were instructed that the language was more grammatically complex than it may have appeared during training, and that they should therefore focus on not just the agreement system, but rather judge any sentence that appeared ill-formed for any reason as ungrammatical.

The testing session consisted of four blocks, each lasting approximately 10 minutes. During the session, participants were seated comfortably in a chair in a sound-attenuated booth. 1500 msec. before the first word of a sentence, a red square appeared in the middle of the screen and remained throughout sentence presentation for fixation purposes. Sentences were once again visually presented above fixation, as they had been during the training session, with each word presented for 300 msec. (500 msec. SOA). Participants read 160 sentences in total: 40 of the fully grammatical sentences (conditions A and D, TABLE), 40 agreement violations (conditions B and E in Table 2.1), 40 transitivity violations (conditions C and F, TABLE 2.1), and 40 auxiliary selection violations (conditions D and G, TABLE 2.1). Across conditions, sentences were divided evenly between transitive and intransitive structures, with each containing an equal proportion of singular and plural agreement relationships. Thus participants saw 120 correct agreement patterns and 40 violations of the agreement pattern that they had been explicitly trained on.

During testing, sentences were pseudo-randomized such that no three sentences from the same condition occurred in succession and no more than six out of ten sentences were from the same condition. Verbs, clause types, and singular/plural agreement relations were balanced across conditions. At the end of each sentence,

participants made grammaticality judgments using response buttons, counterbalanced across response hands. Unlike the training phase of the experiment, participants received no feedback after making their responses. After each judgment was made, a “Press for next” message remained on the screen until the participant chose to proceed to the next stimulus sentence via button press.

2.4 Electrophysiological recording

The electroencephalogram (EEG) was recorded from 26 tin electrodes mounted geodesically in an Electro-cap. These sites included midline prefrontal (MiPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right medial frontal (LMFr and RMFr), left and right dorsal frontal (LDFr and RDFr), left and right medial central (LMCe and RMCe), midline central (MiCe), left and right medial lateral central (LDCe and RDCe), left and right lateral temporal (LLTe and RLTe), left and right medial lateral parietal (LDPa and RDPa), midline parietal (MiPa), left and right lateral occipital (LLOc and RLOc), left and right medial occipital (LMOc and RMOc), and midline occipital (MiOc). Each electrode was referenced on line to the reference electrode at the left mastoid. To monitor blinks and eye movements, electrodes were placed on the outer canthi and under each eye. Impedances were kept below $5K\Omega$ during recording. The EEG was amplified using James Long amplifiers with an online band-pass filter (.01 to 100 Hz), and digitized at a sampling rate of 250 Hz.

2.5 Data analysis

2.5.1 Behavioral training session

The primary goal of the analysis of the behavioral training data was to demonstrate evidence of explicit learning. If there was no evidence of explicit learning within the focus of attention, we would be unable to draw any substantive inferences about implicit learning outside the focus of attention (because it would be essentially impossible to determine what participants were attending to). In order to ascertain that our explicit learning task had been successful, we split the training session into ten blocks of approximately 38 sentences each and looked for evidence of increasing accuracy and decreasing reaction time (RT) across the experimental session, generally assumed in the literature to be evidence that learning has taken place (see Nissen & Bullemer, 1987). In order to verify the statistical significance of these patterns, we conducted trend analyses to determine if there were linear increases and decreases in accuracy and RT, respectively.

2.5.2 Testing session: Grammaticality judgment accuracy

We used two-tailed t-tests to analyze the grammaticality judgment data from the testing session, comparing each condition to all others. Differences in accuracy across conditions would reveal evidence of the transfer of the output of explicit learning from the training phase to the testing phase, as well as potential evidence of implicit learning in unattended conditions outside the focus of attention during training.

2.5.3 Testing session: ERPs

As described below, the main analysis focused on the main verb position, as it was at this position that the critical experimental manipulation was encoded.¹⁶ As we were unsure based on the literature what type of ERP results to expect, and suspected that violations of either the attended and/or unattended grammatical systems might elicit ERP effects typically associated with attentional and/or linguistic processing, we conducted analyses in the N100 (50-150 msec.), P200 (150-250 msec.), LAN/N400 (300-500), and P600 (500-800 msec.) time windows. Mean amplitude measurements were taken over 1000 msec. epochs (including a 100 msec. pre-stimulus baseline), time-locked to the onset of the critical main verb.

Because of the design of the materials, a full factorial ANOVA was deemed inappropriate.¹⁷ However, we did conduct ANOVAs with generic four-level condition factors. Though main effects of – or distributional interactions with – the condition factor were uninterpretable using this approach, we nonetheless pursued this type of analysis because when conducting ANOVAs our statistical software simultaneously conducts post-hoc, Bonferroni corrected pairwise comparisons (family wise error rate = .05; Bonferroni probability = .0083), such that all three violation types could be compared to grammatical sentences (and to each other) in an omnibus ANOVA. This will be referred to as the full analysis. In addition, a subsequent distributional analysis was conducted for each of the three main comparisons, including condition (two levels),

¹⁶ Even in the case of the auxiliary violations, it was only once the main verb was presented that it was possible to determine the grammaticality of the auxiliary-main verb combination.

¹⁷ Due to the fact that we were comparing a single grammatical condition against three different violation types – one (agreement) attended during training and two (transitivity and auxiliary selection) unattended, our materials did not constitute a fully crossed design.

hemisphere (left vs. right), anteriority (prefrontal vs. frontal vs. temporal vs. occipital), and laterality (lateral vs. medial). Electrodes included in the distributional analysis were left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right lateral temporal (LLTe and RLTe), left and right dorsal lateral parietal (LDPa and RDPa), left and right lateral occipital (LLOc and RLOc), and left and right medial occipital (LMOc and RMOc). This distributional analysis was conducted in order to investigate further those effects that appeared to be spatially localized. We felt that this type of analysis was warranted because one of the stated goals of the experiment was to examine the ways in which the brain's responses to each violation type patterned together or differently, and one dimension along which these effects may differ is in their topographies over the scalp. Furthermore, when it was necessary to corroborate local effects suggested by the distributional analysis, ANOVAs were performed at individual levels of these factors. The Huynh-Feldt (1976) correction for lack of sphericity was applied, and corrected p -values are reported with the original degrees of freedom.

2.6 Predictions

Due to the novelty of our experimental paradigm and a shortage of research designed to test our specific hypotheses, our predictions were somewhat limited at the outset of data collection. If selective attention to a grammatical system is necessary for learning to take place, we predicted that we would observe differences in brain and/or behavioral responses during the testing session only for the attended agreement system, and therefore differences would be observed between only the grammatical sentences and agreement violations. If this pattern obtain, it would be consistent with the

literature on the role of selective attention of non-linguistic information. Alternatively, if selective attention to a grammatical pattern is in fact not necessary for learning to take place, we predicted that brain and/or behavioral responses would differ for all violation types when compared to grammatical controls, perhaps differing in onset or amplitude as a function of violation type. Secondly, as reported in McLaughlin et al. (2004) and Tokowicz & MacWhinney (2005), we left open the possibility that we might observe dissociations between brain and behavioral responses during testing, such that responses to violations might be evident in one dependent measure but not the other.

3. Results

3.1 Behavioral data: Training phase

Analysis of the behavioral data revealed a pattern of both increasing accuracy and decreasing RT as the training session progressed. Mean accuracy across participants was at chance levels during the first block, but began to differ significantly from chance during the second block after approximately forty trials of the training paradigm ($t(19) = 2.68$, $p = .008$), with an increase in mean accuracy of 9.6%, indicating that subjects were beginning to acquire aspects of the agreement system very rapidly within a period of ten to twenty trials. Additionally, a trend analysis revealed a robust ($R^2 = .901$, $p < .0001$) positive linear trend of increasing accuracy across all ten blocks of the experiment (see FIGURE 2.1). Mean accuracy for the entire training session was 70.8%, and accuracy during the final block reached 79.8%.¹⁸

¹⁸ In fact ten participants performed, as described below, at between 85% and 100% accuracy during the final block, evidence of extremely robust learning after very brief exposure to the artificial grammar.

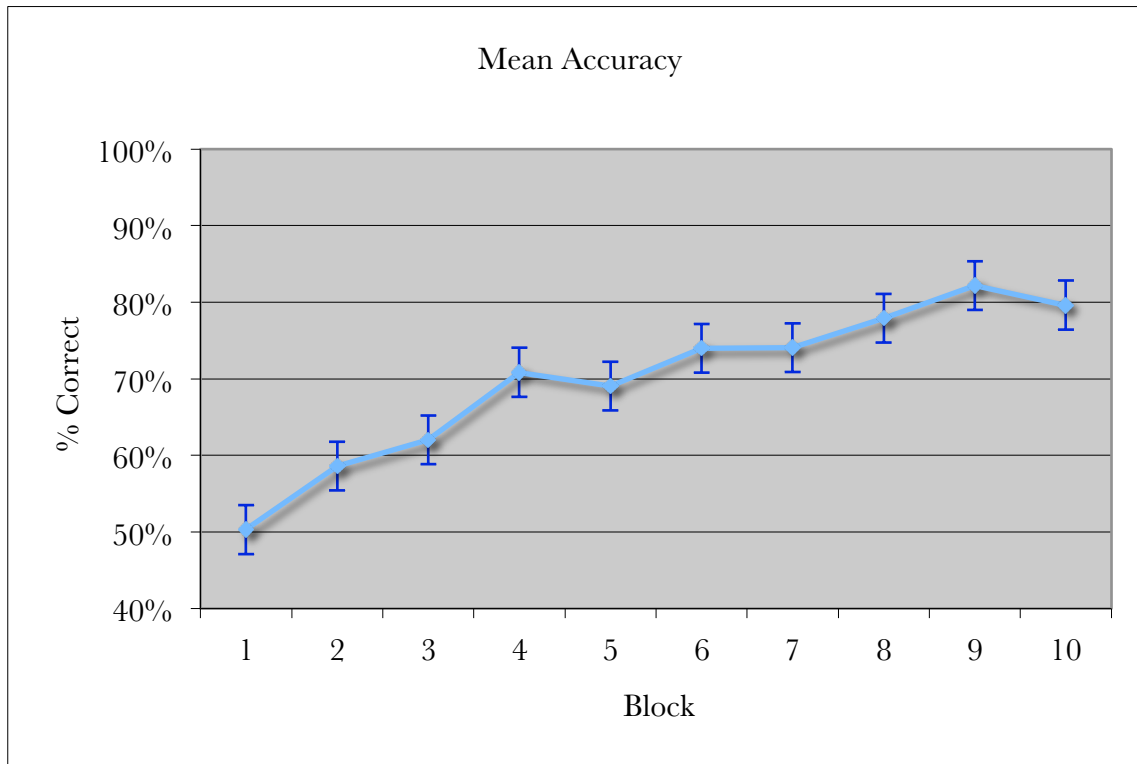


FIGURE 2.1: 2AFC accuracy over time during training session (all participants)

As was the case with the accuracy data, the RT data were also consistent with the occurrence of rapid explicit learning, with a significant difference again observed between the first and second blocks ($t(19) = 3.42, p = .0062$) and a mean decrease in response time of 708 msec. across the entire training session. A trend analysis revealed a strong ($R^2 = .856, p < .0001$) negative linear trend of decreasing RT across all ten blocks of the experiment (see FIGURE 2.2). Based on these results, we felt that we had obtained robust evidence of explicit learning during our (brief) training paradigm, and therefore concluded that our manipulation of attention had been successful.

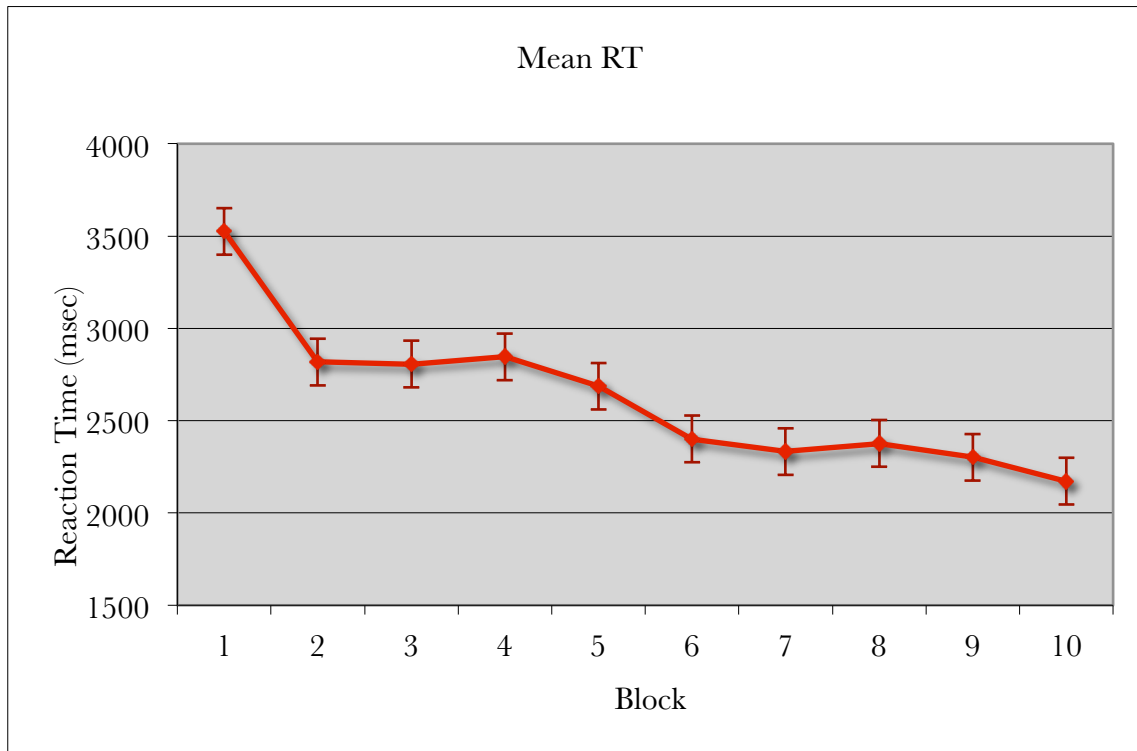


FIGURE 2.2: 2AFC RT over time during training session (all participants)

3.2 Behavioral data: Testing phase

Analysis of the grammaticality judgment data during testing revealed that the accuracy of responses to grammatical sentences did not differ significantly from the accuracy of identifying agreement violations ($t(19) = 1.61$, $p = .177$), but did differ significantly from the accuracy of correctly identifying both transitivity ($t(19) = 3.67$, $p < .01$), and auxiliary selection ($t(19) = 4.61$, $p < .01$), violations. Accuracy in responding to transitivity and auxiliary selection violations did not differ significantly ($p > .15$). It is worth noting that accuracy for both of these conditions was actually below chance ($d' = -1.4$ for auxiliary selection violations and $d' = -.88$ for transitivity violations) suggesting that participants were focusing solely on the agreement system during testing (or had essentially mis-learned these other systems, though we find this

possibility far less likely). Under this scenario, it is not surprising that participants misjudged the transitivity and auxiliary selection violations as grammatical at rates of 76% and 67%, respectively, as these conditions fully conformed to the agreement system that participants had received explicit instruction on during the training session. Based on these results, we believe it is safe to conclude that behaviorally, participants showed evidence of having (explicitly) learned the attended grammatical system, as they were able to judge both grammatical sentences and agreement violations, but showed no evidence of having (implicitly) learned the unattended transitivity and auxiliary selection systems (see FIGURE 2.3).

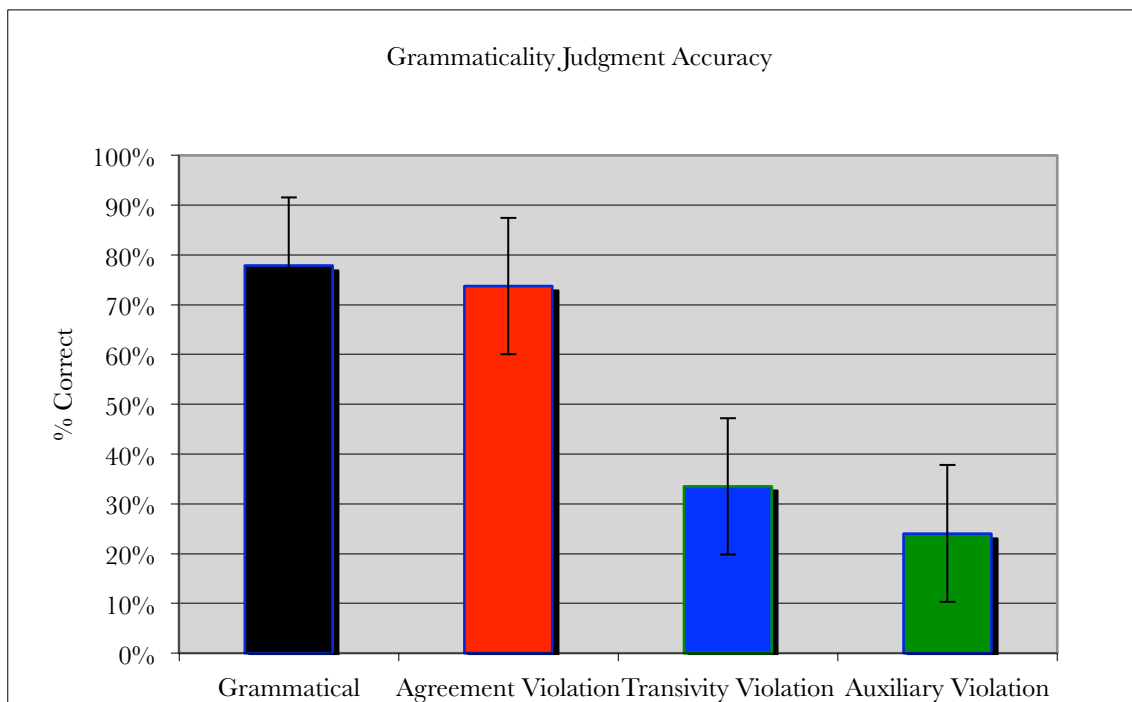


FIGURE 2.3: Grammaticality judgment accuracy in the testing session

3.2 ERP testing data

3.2.1 N100 time window (50-150 msec.)

The results of post-hoc pairwise comparisons showed that grammatical sentences elicited larger N100 responses than auxiliary selection ($p = .0082$) and

transitivity ($p < .0001$) violations but not agreement violations ($p = .279$). Agreement violations also elicited larger N100 responses than transitivity ($p = .0002$) but not auxiliary violations ($p = .071$), though this difference was statistically marginal. Lastly, there were no significant differences between the auxiliary and transitivity violations ($p = .124$). The distributional analysis revealed no meaningful differences (for ERP effects in all time windows measured see FIGURE 2.4.).

3.2.2 P200 time window (150-250 msec.)

In the 150-250 msec. time window, there were again significant differences between conditions in the post-hoc pairwise comparisons. Results of these analyses showed that the responses to grammatical sentences in this time window differed from the responses to all other conditions, with agreement violations ($p = .0012$), transitivity violations ($p = .0016$), and auxiliary selection violations ($p < .0001$) with all violation conditions eliciting more positive responses than grammatical sentences. On the other hand, ERP responses did not differ across violation types (all $p = \text{n.s.}$). The distributional analysis again revealed no meaningful differences.

3.2.3 LAN/N400 time window (300-500 msec.)

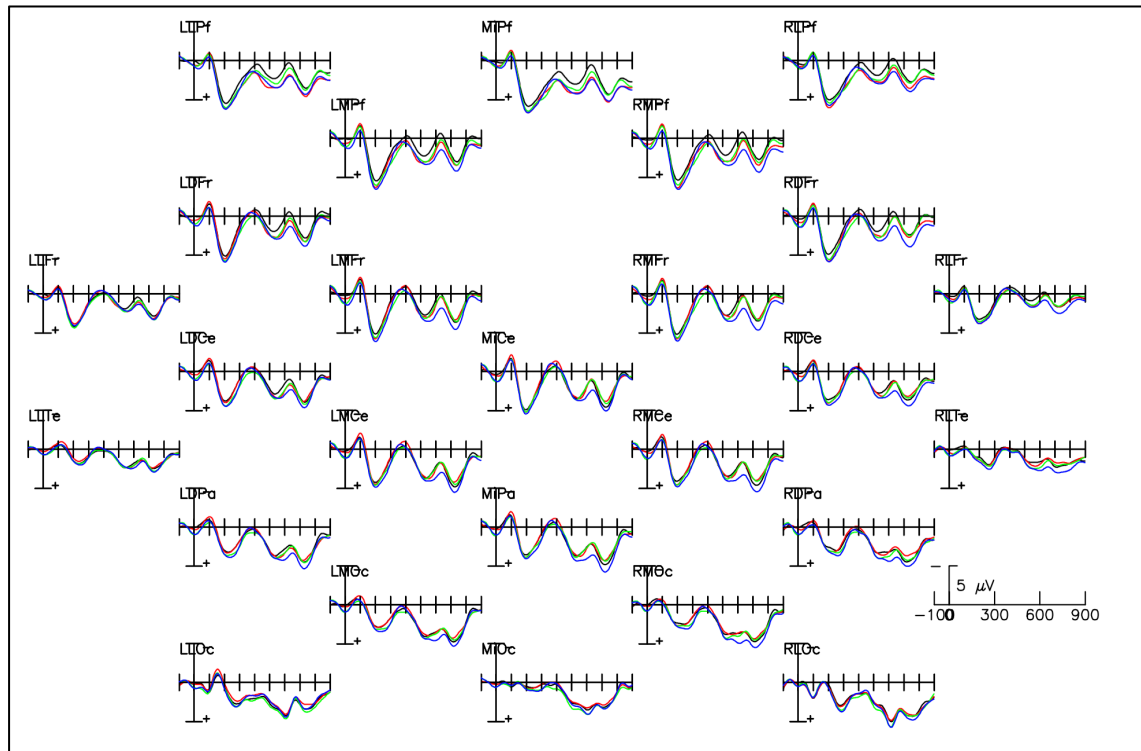
In the 300-500 msec. time window, there were once again significant differences between conditions in the post-hoc pairwise comparisons. The results of these analyses clearly showed that grammatical sentences still differed from agreement, transitivity, and auxiliary selection violations ($p < .0001$ in all cases). No significant differences were found when comparing violation conditions (all $p = \text{n.s.}$). In addition, the distributional analysis revealed that grammatical sentences differed from all violation

types at prefrontal ($p = .00048$) and frontal ($p = .00051$) electrode arrays only (consistent with the frontal distribution indicated in FIGURE 2.5).

3.2.4 P600 time window (500-800 msec.)

In the 500-800 msec. time window, we again observed robust differences between conditions in the post-hoc pairwise comparisons. Results of these analyses showed that the transitivity violations differed from all other conditions ($p < .0015$ in all cases). The distributional analysis revealed no meaningful differences.

(A)



(B)

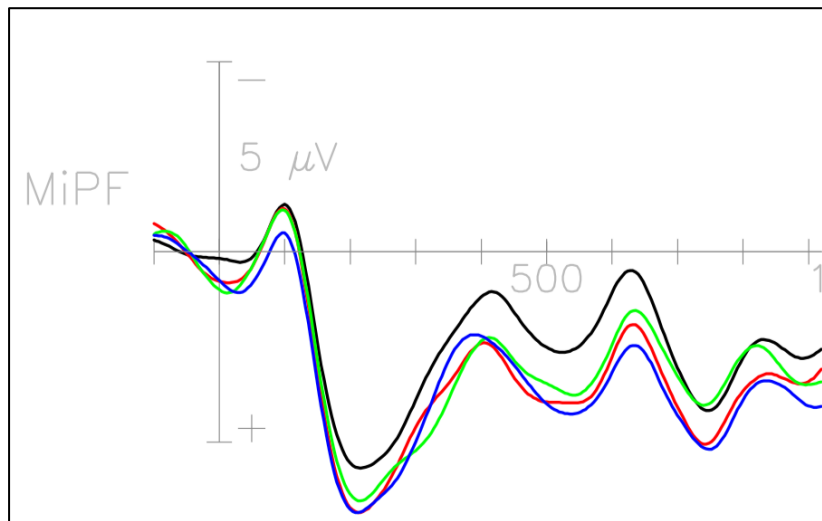
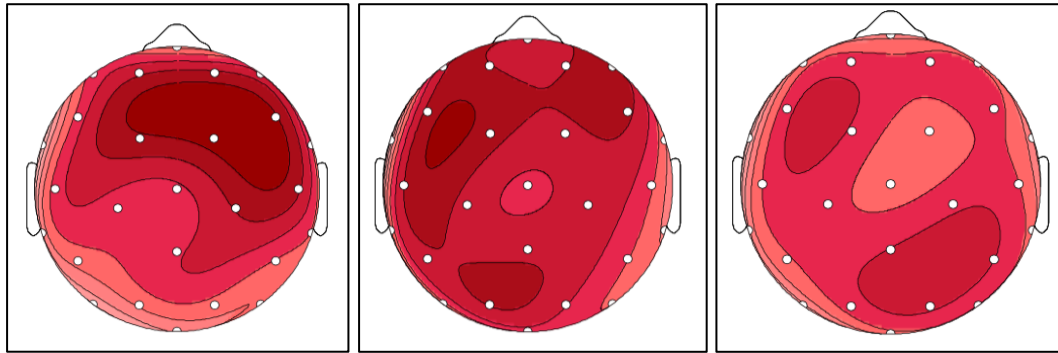


FIGURE 2.4: (A) Grand average ERP waveforms for the **grammatical controls**, **agreement violations**, **transitivity violations**, and **auxiliary selection violations**, for a one-second epoch with 100 msec. baseline. (B) Grand average ERP waveforms for all conditions at electrode MiPF.

(A: (i, ii, iii))



(B: (i, ii, iii))

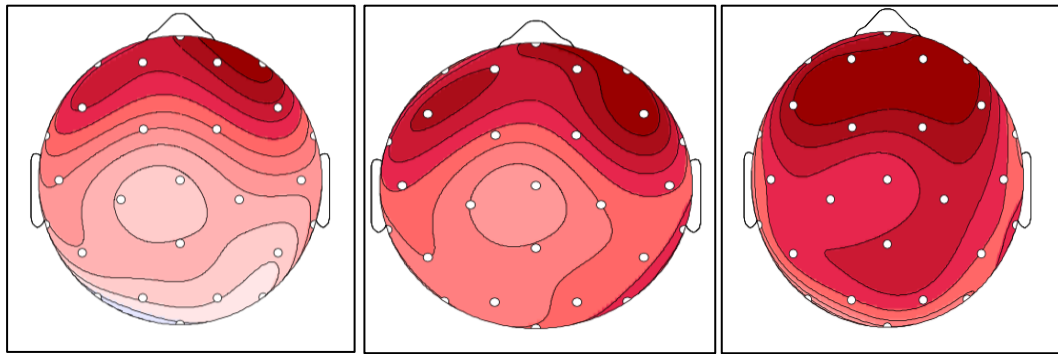


FIGURE 2.5: (A) Topographic scalp isovoltage maps of the mean difference between (i) **grammatical controls** and **agreement violations**, (ii) **grammatical controls** and **transitivity violations**, and (iii) **grammatical controls** and **auxiliary selection violations** in the 150-250 msec. latency window. (B) Identical comparisons for the latency window of 300-500 msec. Microvolt scale standardized for all plots; red indicates increased positivity in response to all violation types compared to control sentences.

3.3 Correlational analyses

3.3.1 Correlations between training performance and behavioral testing accuracy

In analyzing the data from both the training phase and the grammaticality judgment data collected during the testing phase, we noted a great deal of variability in performance across participants. During the training phase, participants' mean accuracy was 70.8% (S.D. = 12.1%; range = 38.5%). During testing, in which participants' knowledge of both attended and unattended grammatical systems was assessed, variability in performance was even more pronounced. For the grammatical controls

mean accuracy was 77.8% (S.D. = 12.4; range = 60.1%), for agreement violations 73.7% (S.D. = 20.3; range = 97.5%), for transitivity violations 33.5% (S.D. = 24.5; range = 67.5%), and for auxiliary selection violations 24% (S.D. = 12.6%; range = 23.2%).

Due to this unexpectedly high degree of cross-participant variability, we conducted unplanned correlational analyses in order to determine whether participants' mean accuracy during training correlated with grammaticality judgment accuracy in each of the four experimental conditions during testing (grammatical controls and agreement, transitivity, and auxiliary selection violations). There were significant correlations in each case. Accuracy on the explicit learning task during training, during which participants were focused solely on the agreement system, correlated *positively* with judgment accuracy in response to grammatical controls ($r(18) = .77$, $p < .0001$) and agreement violations ($r(18) = .74$, $p < .0001$) during testing, but *negatively* with accuracy in judging transitivity ($r(18) = -.453$, $p = .047$), and auxiliary selection violations ($r(18) = -.516$, $p = .039$) during testing (see FIGURES 2.6 – 2.9). Attempted correlational analyses comparing training performance with the amplitude of the ERP effects elicited during testing yielded no significant results, perhaps due to a lack of statistical power.

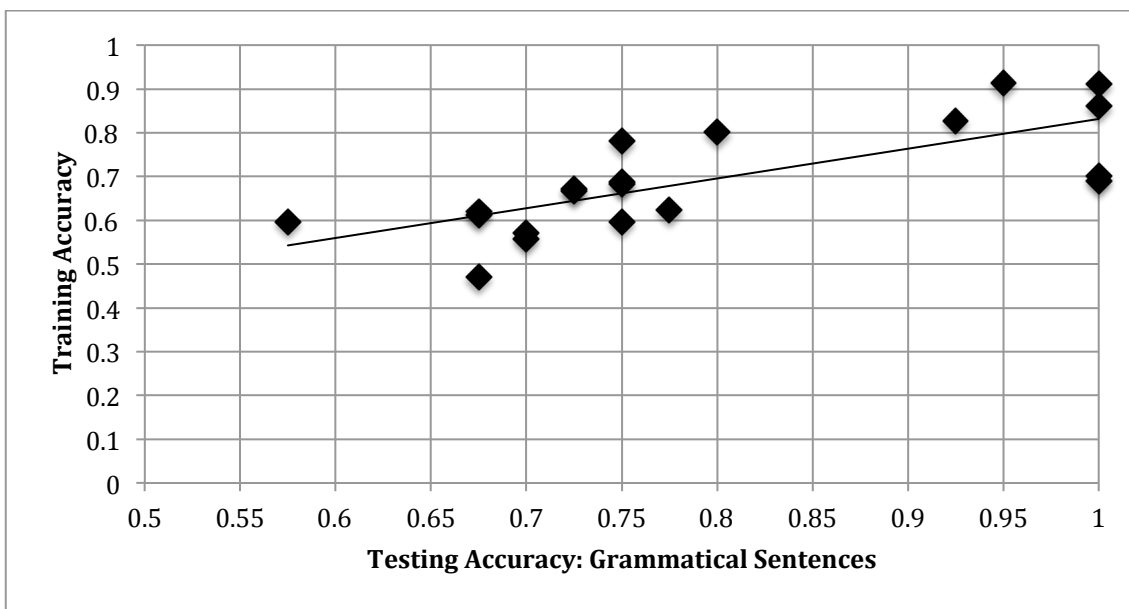


FIGURE 2.6: Correlation between training accuracy and testing accuracy on grammatical sentences

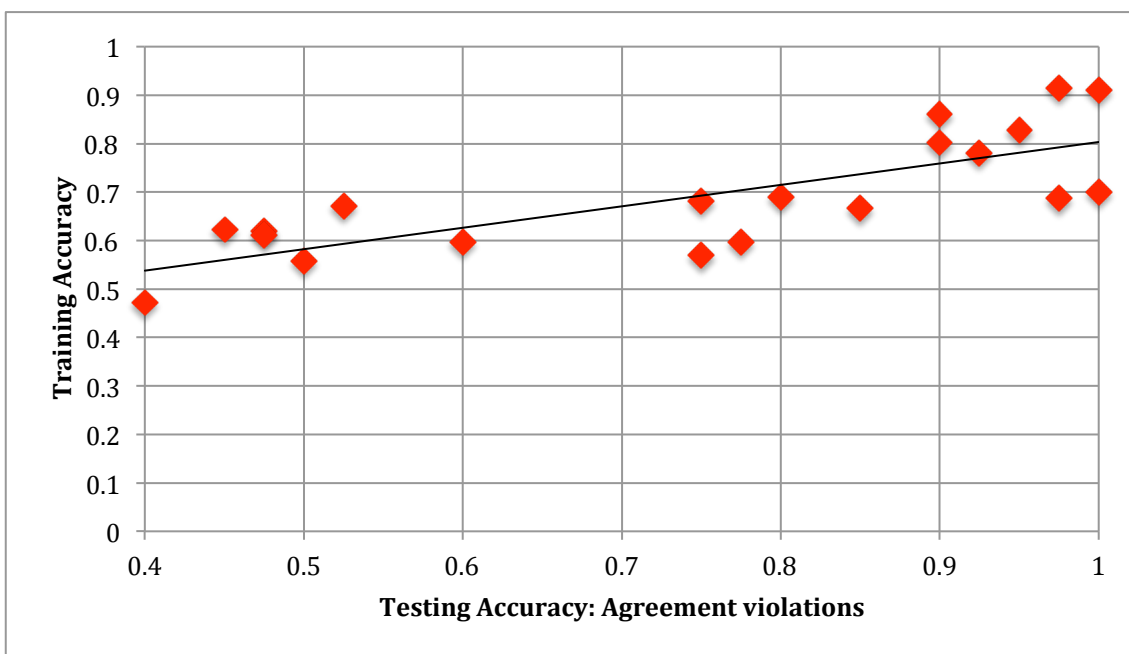


FIGURE 2.7: Correlation between training accuracy and testing accuracy on agreement violations

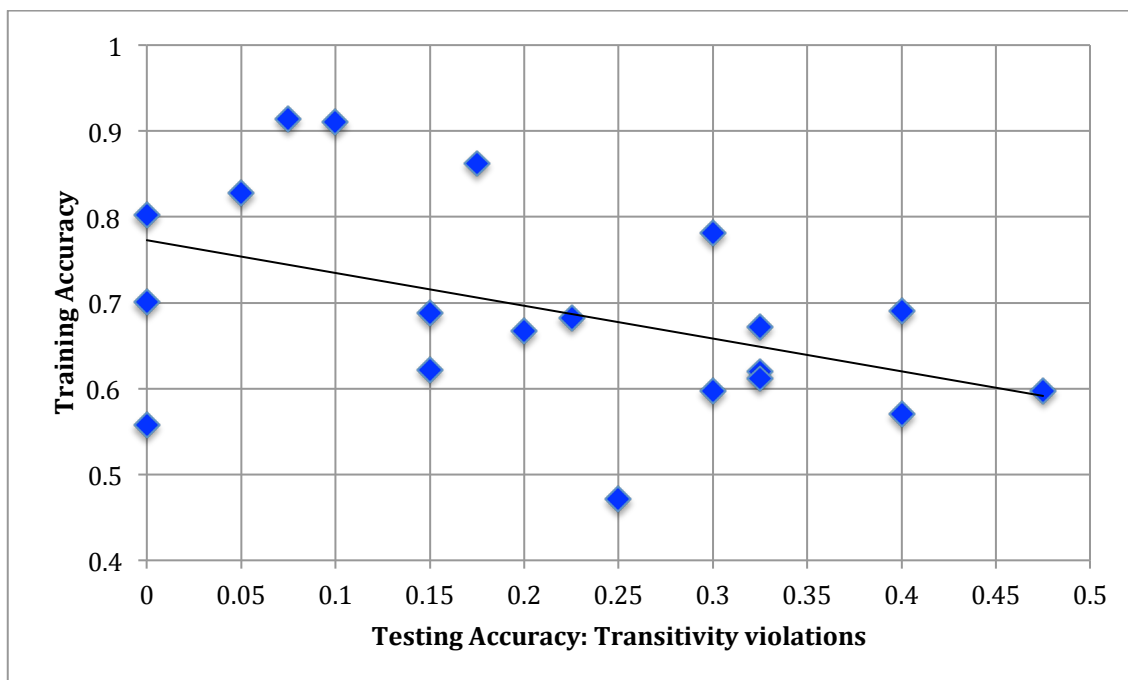


FIGURE 2.8: Correlation between training accuracy and testing accuracy on transitivity violations

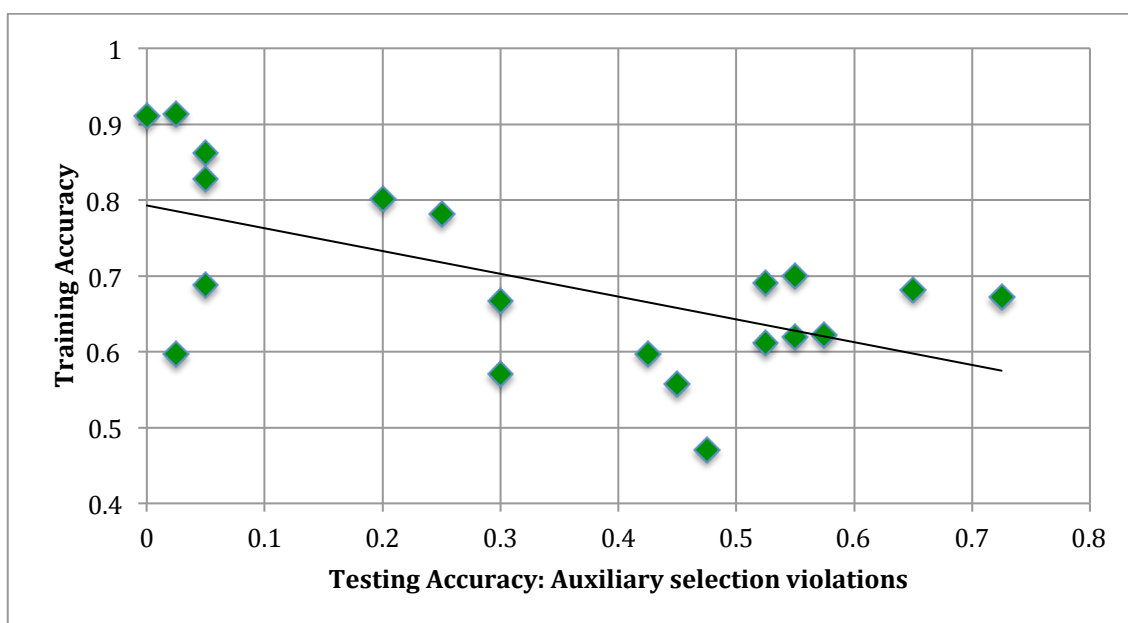


FIGURE 2.9: Correlation between training accuracy and testing accuracy on auxiliary selection violations

This pattern of results suggests, as one would expect, that the more successful participants were at learning the agreement system, the better they were at correctly judging grammatical sentences and agreement violations during testing. In contrast, successful training performance on the agreement system appears to have led to a decreased ability to identify violations of the transitivity and auxiliary systems of the artificial language during the testing phase. This impression is underscored by the fact that the ten participants who were most successful at correctly identifying grammatical sentences (mean accuracy = 86.5%, $d' = 1.08$) and agreement violations (mean = 90.75%, $d' = 1.49$) during testing averaged only 19.1% accuracy ($d' = -1.8$) when judging transitivity violations and 16.1% accuracy ($d' = -1.87$) when judging auxiliary selection violations during testing – substantially below the respective group averages of 33% and 24%. In comparison, the ten subjects who were least successful at correctly identifying grammatical sentences (mean accuracy = 65.1%, $d' = .87$) and agreement violations (mean accuracy = 51%, $d' = .26$), during testing, identified transitivity violations with 47.1% accuracy ($d' = -.6$) and auxiliary selection violations with 37.5% accuracy ($d' = -.2$) during testing – substantially above the respective averages of the group as a whole (33% and 24%, respectively) and indicative of minimal sensitivity to auxiliary selection and transitivity violations (i.e. still at or below chance).

3.4 Summary of results

3.4.1 Behavioral training data

During the 45-60 minute behavioral training phase, we observed evidence of rapid and robust explicit learning of the attended agreement system. This explicit learning manifested in both rapid increases in response accuracy and decreases in

response time between the first and second blocks of the experiment, as well as in statistically robust linear trends of increasing accuracy and decreasing RT across the entire training set.

3.4.2 Behavioral testing data

During the testing phase of our experiment, both grammatical controls and agreement violations, which had been attended to during training, were judged correctly at a rate significantly greater than chance, with accuracy rates not differing significantly between these two conditions. However, responses to these two conditions differed significantly from those to transitivity and auxiliary selection violations, which were not attended to during training. These conditions were judged at below chance levels, and did not differ from one another. Additionally, results of correlational analyses revealed highly significant correlations between training and testing performance, with participants who were most successful during the training session being the least accurate when judging violations of the unattended transitivity and auxiliary selection systems.

3.4.3 ERP testing data

- 50-150 msec. window

In the N100 time window, there were significant differences between the brain responses to grammatical sentences and transitivity violations, with transitivity violations eliciting greater negativity. In addition, the agreement violations differed significantly from the transitivity violations, with the agreement violations eliciting larger amplitude N100 effects. There was also a statistically marginal difference between the responses to grammatical controls and agreement violations.

- 150-250 msec. and 300-500 msec. time windows

In these two time windows, the responses to all violation types were more positive than the responses to grammatical controls but did not differ from each other. In the 300-500 msec. window, these differences were spatially localized to prefrontal and frontal sites, suggesting a more frontal distribution than in the P200 time window.

- 500-800 msec. time window

In this time window there were significant differences only between the transitivity violations and the three other conditions.

4. Discussion

4.1 The role of attention in implicit language learning

As stated in the introduction, the role of attention in implicit learning has thus far been incompletely described, especially in the case of implicit language learning. Manipulations of selective attention appear to disrupt implicit learning and its behavioral outputs consistently in non-language learning contexts (Jimenez & Mendez, 1999; Jiang & Chun, 2001; Hoffman & Sebold, 2005). However, manipulations of attentional load have been claimed both to influence (Nissen & Bullemer, 1987; Toro et al., 2005) and not to influence (Saffran et al., 1997) learning in both linguistic and non-linguistic tasks. The present study focused on the role of selective attention in implicit learning language learning, using a manipulation that has no precedent in the literature. The current study was designed to build on the non-linguistic literature on implicit learning. We adopted a selective attention paradigm that manipulated participants' attention during a language-learning task by directing their attention to only one set of regularities present in the training input. Our results showed that this manipulation of

selective attention disrupts implicit language learning (in a graded fashion), though participants' brain responses indicated that they had extracted information about the unattended regularities in the training materials. It should be noted, however, that our results may not generalize to *all* linguistic regularities, as the current study focused solely on unattended patterns of verbal subcategorization. The extent to which the same pattern of results would obtain if we examined the role of attention in the learning of other types of unattended linguistic regularities is an open, empirical question. Therefore, our data are to some extent consistent with the literature of implicit learning of non-linguistic information (Jimenez & Mendez, 1999; Jiang & Chun, 2001; Hoffman & Seibald, 2005), in that they show that selection attention to a set of regularities is necessary for evidence of implicit learning to be apparent in behavioral measures. However, the pattern of brain responses observed suggest that *some* form of implicit learning is taking place, and the extent to which the same pattern would obtain in studies of implicit learning of non-linguistic information is an open, empirical question. The two dependent measures that we used in this study, overt behavioral and covert brain responses (referred to as "explicit vs. implicit processing" in Tokowicz & MacWhinney 2005), point to two seemingly opposite conclusions about the role of selective attention in adult language learning. While we think that these conclusions are ultimately compatible, a claim to which we will return in the end, initially we will discuss the consequences of each dependent measure separately.

4.2 Behavioral responses

During testing, participants were below chance at identifying violations of the grammatical systems that were unattended during training (i.e. transitivity and auxiliary

selection violations). This would seem to support the hypothesis that selective attention is necessary for successful implicit (language) learning. Put differently, it appears to be the case that when a language learner's attention is directed away from a statistical regularity or co-occurrence restriction in language input, implicit learning is blocked (at least as assessed by binary grammatical judgments, the extent to which we would observe the same pattern with a more graded behavioral measure remains a topic for future research).

Nevertheless, within this general overall failure to acquire unattended grammatical patterns, our correlational analysis suggests that blockage of implicit learning occurred in a graded rather than discrete fashion. Those participants who were best able to learn the attended agreement system during training performed worst during testing on judgments of the unattended transitivity and auxiliary selection violations, while those participants who were less successful at learning the agreement system were better able to detect transitivity and auxiliary selection violations during testing – although it must be emphasized that these participants were still at or below chance on the grammaticality judgment task (and as such, to be clear, we are not making the claim that these participants were actually learning the grammatical systems unattended during training). Regardless of this across-the-board poor performance when judging the grammaticality of the grammatical systems unattended during training, the positive training-testing correlations seen in the “good learners” on judgments of grammatical controls and agreement violations are highly consistent with the notion that the influence of selective attention on implicit learning is graded rather than discrete.

As stated above, it is worth noting that accuracy for both of these conditions was approximately at chance (see FIGURE 2.3), suggesting that participants were focusing solely on the agreement system during testing (as the d' scores indicate this is most likely due to a strong “yes, it’s grammatical” bias for the auxiliary selection and transitivity violations). Under this scenario, it is not surprising that participants misjudged the transitivity and auxiliary selection violations as grammatical at rates of 76% and 67%, respectively, as these conditions fully conformed to the agreement system that participants had received explicit instruction on during the training session. Under the assumption that subjects were judging these violations as grammatical because the agreement relations they contained were well-formed, we conducted t-tests that showed that accuracy in responding to transitivity (67%) and auxiliary selection violations (76%) differed from chance but did not differ significantly from accuracy on judgments of grammatical controls and agreement violations, the exact that pattern that one would predict if participants had been focusing solely on the agreement system during training.

A straightforward interpretation of these results is that poor learners were not as successful at selectively attending the agreement system they were supposed to be learning. Again, these results are consistent with the training-testing correlations observed, and provide strong evidence in favor of the position that selective attention plays a crucial role in implicit learning of linguistic input.

4.3 Brain responses

In contrast to the behavioral responses, the brain responses elicited in our study indicated sensitivity to all violation types. We observed a complex pattern of responses

in the N100 time window that, though somewhat difficult to interpret, nonetheless distinguished between grammatical sentences and violations. The pattern of results in the 150 – 250 msec. and 300 – 500 msec. time windows was clear: grammatical sentences differed from all violation types, which did not differ from one another. Lastly, between 500 – 800 msec. there were significant differences between transitivity violations and all other conditions.

These results support the hypothesis that selective attention is ***not*** in fact necessary for implicit learning to output at least some form of learned representation at the neural level, even absent behavioral evidence of this learning. At some level of representation in the brain, regularities governing all three patterns had been (whether consciously or unconsciously) extracted from the data that participants were exposed to during training – which, it should be reemphasized, lasted only about 45 to 60 minutes. This suggests that, even though overt behavioral responses may show no evidence of it, representations of the statistical regularities governing the input are formed in the brain whether such regularities are attended to or not. However, the fact that the components that were modulated, the N100, P200, and (arguably) a P300 effect (see discussion below) onsetting between 300 – 500 msec. and continuing on into the 500-800 msec. time window, are typically interpreted as indexing domain-general cognitive operations suggests that at this early stage of learning, when language proficiency is low, the language input is simply being interpreted as a set of systematic regularities, the processing of which has yet to “passed on” to the language system for processing.¹⁹

¹⁹ See Liégeois et al. (2003), Leonard et al. (2010) and Mayberry et al. (2011) for similar arguments.

We want to emphasize that the exact nature of the brain responses elicited by violations of the artificial language's grammar during testing is not crucial to the larger point we wish to make here, which is that selective attention does play a role in implicit language learning. Our preferred interpretation, consistent with the data, is that the early exogenous effects between 50 – 150 msec. index selective attentional modulations of the domain-general N100 (Hillyard, 1973). Grammatical sentences and agreement violations elicited larger N100s compared to transitivity and auxiliary selection violations, which did not differ from one another. We believe that these increases in N100 amplitude index are a result of the fact that the grammatical sentences and agreement violations contain sequences that the participants were exposed to during training. Consequently, these task-relevant sequences were selectively attended leading to increases in N100 amplitude.

Subsequent to these attentional modulations of the N100, there were larger P200 responses to all violation types compared to grammatical controls (see Luck, 2014 for a brief review of the cognitive processes that the P200 is argued to index and Furutsuka, 1989 for a discussion of this component in the context of attention). This increase in P200 amplitude plausibly indexes processes of feature-based selection of meaningful (but unexpected) stimuli, i.e. violations. Alternatively, this effect could reflect additional attentional processes occurring further downstream from those occurring in the N100 time window. Under this account, agreement violations capture attention due to their behavioral relevance and the unattended violations because they constitute unfamiliar sequences that participants have not encountered before.

With regards to the effects observed in the 300 – 500 and 500 – 800 msec. time windows, there are three interpretive possibilities for a positivity in this latency range: a P3a, a P3b, or a P600. If this effect were to be interpreted as a P3a, it would plausibly index the processing of novelty (Polich, 2007), as each violation condition is *relatively* unfamiliar/novel to participants. Though the apparent frontal distribution of the effect between 300 – 500 msec. is consistent with this interpretation (though the distributional analysis did not yield any significant condition x anteriority interactions), we do not believe that this interpretation best accounts for the data, as the violation conditions were not truly novel to the extent that the deviant stimuli in novelty oddball paradigms that elicit P3a responses typically are (see Debener et al., (2005) for a description of the novelty oddball paradigm). The stimuli that elicit P3a effects in these oddball paradigms are very low frequency and highly deviant (for example, an environmental sound occurring with ten percent probability in a series of high and low tones). In contrast, our materials contained 40 items per condition (with each violation occurring with 25% probability), and the violations were therefore not novel or unexpected to the same extent. Additionally, the P3a-eliciting stimuli in oddball paradigms are task-irrelevant, while the main verbs eliciting these effects in the current study were crucial to task performance. As such, though the latency and distribution of these effects is consistent a P3a interpretation, we do not feel that such an interpretation is merited based on our experimental design and the nature of the behavioral task employed.

If this effect is a P600 indexing parsing difficulty (Osterhout & Holcomb, 1992), then it is difficult to provide an explanation of why only the transitivity violations differed from the grammatical controls in the 500 – 800 msec. time window, while all

three-violation conditions differed from controls in the preceding time window. Though arguments have been made that positivities with different distributions and onset latencies index different late syntactic processes of revision, repair, and re-analysis (for example, see Friederici et al., 2002 and Kaan & Swaab 2003), we are hesitant to infer strong mappings between our violation types and which of these syntactic processes they may engage. As such, we feel that interpreting the positivities seen in these time windows as P600 effects would be purely speculative.

We therefore feel that the interpretation that is most consistent with the data is that the positivities observed between 300 – 500 and 500 – 800 msec. are actually a P3b effect potentially indexing processes of event categorization (Kok, 2001), in this case reflecting processes of categorizing sentences as grammatical or ungrammatical. Under this view, the fact that it was the transitivity violations that elicited enhanced positivity between 500 – 800 msec. may result from the fact that these violations are most difficult to categorize as ungrammatical due to the fact that participants must take the entire sentence into account in order to do so. Once more, regardless of how these effects are interpreted, it is important to note that we do not feel that the functional identity of the observed responses is crucial to our claim that selective attention influences implicit learning.

4.4 Brain-behavior dissociations and the output of implicit learning

Both MacLaughlin et al. (2004) and Tokowicz & MacWhinney (2005) reported patterns of brain/behavior dissociation after extensive explicit language instruction: in each case, evidence of learning was apparent in brain responses (N400 effects in the former case and P600 effects in the latter) but largely absent in participant's behavioral

responses. Here we extend these findings, albeit in the form of modulations of different (domain-general) components, by observing the same pattern in an implicit learning task using a type of attentional manipulation heretofore absent from the language-learning literature. Though subjects showed no behavioral evidence of having learned grammatical patterns unattended during training, their brain responses nonetheless indicated that they had extracted and were sensitive to the regularities encoded in the (unattended) input. Taken together, the results of these ERP studies suggest that the representations formed during early stages of learning (whether explicit or implicit) may be too unstable to influence behavioral outcomes, and only with increased exposure do these representations reach a significantly robust level of activation and/or stability to exert such a behavioral influence. Based on our results, the amount of input necessary to create representations that are stable enough to be detectable in brain responses is minimal, even if such brief input is insufficient to influence behavioral outcomes, a finding that should come of no surprise to adults attempting to learn a second language.

4.4 Consequences for adult second language acquisition

There has long been a distinction made in the theoretical second language acquisition literature between explicit and implicit learning (Bialystock & Fröhlich 1972), or between “learning” and “acquisition” (Krashen 1982). Krashen defines learning as conscious knowledge of a second language: knowing rules, and being both aware of them and able to talk about them. Learning is therefore posited to have only an after-the-fact editing function that can alter the form of an utterance after it has already been produced, either in speaking or in writing. Acquisition, on the other hand, is a

subconscious process, much akin to the way in which children acquire their first language and produce it without thinking about it.

While this distinction is no longer itself an active topic of discussion in second language pedagogy, it seems to us to be a useful one to appeal to in making sense of the results of our study. Participants clearly “learned” the agreement system of the artificial language we created in an entirely different way from the way in which they “learned” the transitivity and auxiliary selection systems (though see section 4.3). Learning of the agreement system was entirely conscious: during training, participants were told explicitly to focus on learning the agreement system, shown examples of subject-verb agreement in English to orient them, given a binary forced choice of two possible agreement patterns for each of nearly 400 sentences that they saw, and given explicit feedback on each sentence as to whether the choice that they had made was correct or incorrect (though they were never informed as to why).

In contrast, no mention whatsoever was made of the transitivity and auxiliary selection systems that were also encoded in the verb phrase of our stimulus materials. While we have no absolute guarantee that participants did not on some level become aware of the transitivity and auxiliary selection patterns in the data they saw, there are three countervailing reasons why we think that they did not. First, the auxiliary of the verb phrase was always presented along with the main verb (which had an agreement-marking prefix) in the binary forced-choice task. While the auxiliary was thus the same across the two choices participants saw (which differed only in the agreement-marking prefix), one could possibly argue that this mode of presentation could have inadvertently served to focus participants’ attention on main and auxiliary verb pairings,

i.e. on the auxiliary selection system, in which case “learning” of auxiliary selection would not be entirely unconscious. However, even if this were true, the same explanation cannot be given for the unconscious “learning” of the transitivity system, which could only be read off the entire sentence structure. But stimulus sentences were never presented in their entirety during training (because we used RSVP for both training and testing) unless participants responded with the wrong agreement pattern choice, in which case the whole sentence did reappear on the screen without further comment. Yet the brain responses to transitivity and auxiliary selection violations were effectively the same between 150 and 500 msec. Furthermore, participants indicated that they had been unaware of the transitivity and auxiliary selection patterns in the data in their post-experiment debriefing questionnaires. And finally, of course, behavioral performance on detecting errors was below chance for both transitivity and auxiliary selection violations.

We therefore think it safe to claim that participants explicitly learned the agreement system, as evidenced by their behavioral performance on the grammaticality judgment task during testing, but implicitly learned (or “acquired”) the transitivity and auxiliary selection systems along with the agreement system, as evidenced in their brain responses to all three types of violations. Intriguingly, this suggests that the outcomes of short-term explicit vs. implicit learning may not be the same: intensive short-term explicit learning seems to result in behaviorally detectable outcomes, at least in the form of the ability to assess the well-formedness of utterances, but short-term implicit learning appears to be sufficient to establish the brain representations that will ultimately serve as the basis for forming such judgments, on the assumption that all

competent users of a language over time establish some set of criteria, however imperfect, for what constitutes an acceptable utterance in the language they use.

This view of things is strikingly reminiscent of a view expressed by Hockett (1948):

The essential difference between the [analytical] process in the child and the procedure of the linguist is this: the linguist has to make his analysis overtly, in communicable form, in the shape of a set of statements which can be understood by any properly trained person, who in turn can predict utterances not yet observed with the same degree of accuracy as can the original analyst. The child's 'analysis' consists, on the other hand, of a mass of varying synaptic potentials in his central nervous system. The child in time comes to *behave* the language; the linguist must come to *state* it.

For Hockett, the outcome of language acquisition is thus “a state of affairs in the nervous system rather than a set of statements.” To the extent that “a set of statements” about the language is tantamount to Krashen’s (1982) “conscious knowledge of a second language, knowing the rules, being aware of them, and being able to talk about them,” Hockett’s distinction appears to foreshadow the explicit vs. implicit learning distinction and map equally transparently onto the pattern of results that we obtained.

At the same time, it must be pointed out that long-term, less intensive explicit learning is no guarantee of behaviorally detectable outcomes: both McLaughlin et al. (2004) and Tokowicz & MacWhinney (2005) report dissociations between the brain and behavioral responses of long(er)-term adult learners of natural second languages. Though the authors do not directly report it, one can only assume that vocabulary acquisition is a very explicit learning task in virtually every language learning setting.

Tokowicz & MacWhinney (2005) report even longer-term brain/behavior dissociations in acquiring morphosyntactic patterns of Spanish. However, participants

were essentially at chance at detecting violations of all three types, while their brain responses in the form of a late positivity effectively distinguished violations from correct control sentences. Most strikingly, gender agreement violations were detected at below chance levels (significantly different from detection accuracy on the other two violation types) but elicited the most robust brain responses.

This raises the question of just how effective explicit learning is in the long term as a language acquisition tool. Krashen (1982) claims that explicit learning is useful only when learners are focused on form (Burt & Dulay, 1978) rather than meaning, and have the time necessary to apply their conscious knowledge of rules. In more recent years, the distinction between explicit and implicit learning, and the claims associated with it, have transmogrified in the second language pedagogy literature into a debate over the best way to induce a cognitive state that maximizes the learner's potential. One such competing approach has crystallized out as what is now called "focus on form" (Lightbown & Spada 1990; Spada 1997; Ellis 2001, 2002; Abu Radwan 2005), in which language learners are explicitly instructed to focus on a specific form-based aspect of the language to the exclusion of others. For example, students may be instructed to focus solely on patterns of agreement (as in the current study) while ignoring semantic or other grammatical information, or vice versa. These types of approaches are typically opposed to so-called immersion-based approaches, in which students are simply immersed in a novel language environment with no explicit instruction to attend to any aspect of the input – analogically similar to the way in which we exposed the participants in our own study to transitivity and auxiliary selection violations.

Empirical results have been somewhat contradictory, with various studies finding that each approach leads to more successful learning outcomes. Based on our results, we think that this may merely point to the fact that both approaches are effective in different ways. While it was never our intention to test explicit hypotheses about adult second language instruction, it is obvious that our results nonetheless do have some bearing on this debate, and we would suggest that they cut both ways.

Our training paradigm (unintentionally) adopted a focus on form approach to learning, explicitly instructing our participants to focus exclusively on the agreement system while ignoring other statistical patterns in the data. In the short term, these explicit instructions appeared highly successful: participants learned the agreement system rapidly, in only 45-60 minutes, as evidenced by patterns of increasing accuracy and decreasing response times during training, and by high accuracy on the grammaticality judgment task during testing. Focusing on the form of the agreement markers thus yielded rapid and robust learning. Whether such impressive gains from concentrated, intensive, and focused training would hold over the long term is open to question, however, based on the long-term natural language learning results reported by Tokowicz & MacWhinney (2005), in which students were still at chance at detecting basic gender and number agreement violations in Spanish after as much as two full years of instruction.

At the same time, it cannot be overlooked that we also found evidence that unattended systems had been “acquired” (to some extent) in the form of brain responses elicited during testing. We suggested above that such brain representations could lay the foundation for conscious knowledge of the system that can eventually be

tapped into and surface in behavioral measures. A wide variety of brain imaging research indicates that mere exposure to a language alters cortical anatomy (Mechelli et al. 2004; Stein et al. 2010; Mårtensson et al. 2012; Grogan et al. 2012), and our results suggest for the first time that passive and unconscious exposure to statistical regularities in language-like input alters brain responses in the very short term, after only an hour or so of exposure. And yet it is also clear from our data that focusing on only one aspect of a grammar in the short term does not block the implicit learning – in the form of emerging neural representations – of either attended or unattended forms: brain responses to violations of all three types in our data were indistinguishable from each other but, in some time windows, significantly different from the responses to grammatical sequences. In other words, while it is true that the brain was able to extract the information it needed to detect irregularities in the agreement system after intensive focused training, it was equally able to extract the information it needed to detect irregularities in the transitivity and auxiliary selection systems with no explicit training whatsoever. In view of this fact, it seems to us that language instruction cannot err, let alone do any harm, as long as it is configured in a way that exposes students to language input by any means whatsoever. In this sense, both language teachers and language learners can take comfort in the fact that pretty much all they need in order for effective language acquisition to occur is to show up for class. And paying attention won't hurt either.

Chapter 2, in full, is currently being prepared for submission for publication. Barkley, C., Kluender, R., & Kutas, M (*in preparation*). The dissertation author was the primary investigator and preparer of this manuscript.

5. Works cited

- Abu Radwan, A. (2005). The Effectiveness of Explicit Attention to Form in Language Learning. *System*, **33**, 69-87.
- Baker, C.I, Olson, C., & Behrmann, M. (2004) Role of attention and perceptual grouping in visual statistical learning. *Psychological Science* **15**, 460–466.
- Burt, M., & Dulay, H. (1978). Some Guidelines for the Assessment of Oral Language Proficiency and Dominance. *TESOL Quarterly*, **12**(2), 177-192.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Boston: MIT Press. Print
- Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press. Print
- Cohen, A., Ivry, R., & Keele, S. (1990). Attention and Structure in Sequence Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 17-30.
- Debener, S., Makeig, S., Delorme, A., & Engel, A. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cognitive Brain Research*, **22**, 309-321.
- Dixon, R.M.W. (1994). *Ergativity*. Cambridge University Press.
- Eimas, Siqueland, Jusczyk, & Vigorito (1971). Speech Perception in Infants. *Science*, **171**, 303-306.
- Ellis, N.R., Palmer, R.L., & Reeves, C.L. (1988). Developmental and Intellectual Differences in Frequency Processing. *Developmental Psychology*, **24**, 38-45.
- Ellis, R. (2001). Introduction: Investigating Form-based Instruction. *Language Learning*, **51**, 1-46.
- Ellis, R. (2002). Does Form-based Instruction Affect the Acquisition of Implicit Knowledge. *Studies in Second Language Acquisition*, **24**, 2223-236.
- Friederici, A., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, **99**, 529-534.
- Friederici, A., Hahne, A., & Saddy, D. (2002). Distinct Neurophysiological Patterns Reflecting Aspects of Syntactic complexity and Syntactic Repair. *Journal of Psycholinguistic Research*, **31**, 45-63.
- Furutsuka, (1989). Effects of rapid attention switching on the N1-P2 amplitude of the visual event-related potentials. *Research & Clinical Center for Child Development Annual Report*, **11**, 55-64.
- Gomez, R.L (2002). Variability and Detection of Invariant Structure. *Psychological Science*, **13**, 431-436.
- Grogan, A., Parker Jones, O., Ali, N., Crinion, J., Orabona, S., Mechias, M., Ramsden, S., Green, D., & Price, C. Structural Correlates for Lexical Efficiency and Number of Languages in Non-native Speakers of English. *Neuropsychologia*, **50**, 1347-1352.
- Hasher, L., & Chromiak, W (1977). The Processing of Frequency Information: An Automatic Mechanism? *Journal of Verbal Learning and Verbal Behavior*, **16**, 173-184.

- Hirsh-Pasek, K., & Golinkoff, R.M. "The Intermodal Preferential Looking Paradigm: A Window onto Emerging Language Comprehension." *Methods for Assessing Children's Syntax*. Ed. McDaniel, D., McKee, C, and Cairns, H.S. MIT Press, 1996. pp 105-105. Print.
- Hockett, C. F. 1948. A Note on 'Structure'. *International Journal of American Linguistics*, **14**, 269-271.
- Hoffmann, J., & Sebal, A. (2005). When obvious covariations are not even learned implicitly. *European Journal of Cognitive Psychology* *17*, 449–480.
- Huynh, H, & Feldt, L (1976). Estimation of the Box Correction for Degrees of Freedom From Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational and Behavioral Statistics*, **1**, 69-82.
- Jiang, Y., & Chun, M. M. (2001). Selective Attention Modulates Implicit Learning. *Quarterly Journal of Experimental Psychology*, **54A**, 1105-1124.
- Jimenez, L, & Mendez, C. (1999). Which Attention is Needed for Implicit Sequence Learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 236-259.
- Johnston, W.A., & Dark, J. (1986). Selective Attention. *Annual Review of Psychology*, *37*, 43-75.
- Jusczyk, P., Cutler, A., & Redanz, L. (1993). Infants' sensitivity to predominant stress patterns in English. *Child Development*, *64*, 675-687.
- Jusczyk, P., Houston, D.M, & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159-207.
- Kaa, E., & Swaab, T. (2003). Repair, Revision, and Complexity in Syntactic Analysis: An Electrophysiological Differentiation. *Journal of Cognitive Neuroscience*, **15**, 98-110.
- Kok, A. (2001). On the Utility of P3 Amplitude as a Measure of Processing Capacity. *Psychophysiology*, **38**, 557-577.
- Krashen, S. (1982). Principles and Practices in Second Language Acquisition. *Pergamon Press*.
- Kutas, M., & Baldwin, K. (1997). An ERP Analysis of Implicit Structured Sequence Learning. *Psychophysiology*, **34**, 74-86.
- Leonard MK, Brown TT, Travis KE, Gharapetian L, Hagler DJ, Dale AM, Elman JL, Halgren E. (2010) Spatiotemporal dynamics of bilingual word processing. *Neuroimage*, **49(4)**, 3286-94.
- Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI Abnormalities Associated with *FoxP2* Gene Mutation. *Nature Neuroscience*, **6**, 1230-1237.
- Lightblown, P., & Spada, N. (1990). Focus on Form and Corrective Feedback in Communicative Language Teaching. *Studies in Second Language Acquisition*, **12**, 429-428.
- Luck, S. (2014). An Introduction to the Event-related Potential Technique. Boston: MIT Press. Print
- Mårtensson, J., Eriksson, J., Bodammer, N.C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of Language-related Brain Areas after Foreign Language Learning. *NeuroImage* **63(1)**, 240-244.

- Mayberry, R., Chen, J.-K., Witcher, P., & Klein, D. (2011). Age of acquisition effects on the functional organization of language in the adult brain. *Brain and Language*, **119**, 16-29.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural Correlates of Second Language Word Learning: Minimal Instruction Production Rapid Change. *Nature Neuroscience*, **7**, 703-704.
- McLaughlin, J., Tanner, D., Pitkanen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete States of L2 Grammatical Learning. *Language Learning*, **60**, 123-150.
- Mechelli, A., Crinion, J.T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R.S., Price, C.J. (2004). Structural Plasticity in the Bilingual Brain. *Nature*, **431**, p. 757.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, **29**, 144-178.
- Newport, E. (1988) Constraints on learning and their role in language acquisition: studies of the acquisition of American Sign Language. *Lang. Sci.*, **10** 147-172
- Newport, E., & Aslin, R. (2004). Learning at a distance: 1. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, **48**, 127-162.
- Nissen, M., & Bullemer, P. (1987). Attention Requirements of Learning: Evidence from Performance Measures. *Cognitive psychology*, **19**, 1-32.
- Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning*, **56**, 199-230.
- Osterhout, L., Poliakow, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., & Hirshcensho, J. (2008). *Journal of Neurolinguistics*, **21**, 509-521.
- Nissen, M., & Bullemer, P. (1987). Attention requirements of learning: evidence from performance measures. *Cognitive psychology*, **19**, 1-32.
- Perruchet, P., & Pacton, S. (2006). Implicit Learning and Statistical Learning: One Phenomenon, Two Approaches. *Trends in Cognitive Sciences*, **10(5)**, 233-238.
- Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Journal of Clinical Neurophysiology*, **118**, 2128-2148.
- Reber, A.S. (1967). Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behavior*, **6**, 855-863.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge: an essay on the cognitive unconscious*. New York: Oxford University Press.
- Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8 month old infants. *Science*, **274**, 1926-1928.
- Saffran, J., Newport, E., & Aslin, R. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, **35**, 606-621.

- Saffran, J., Newport, E., Aslin, R., Tunick, R., & Barruceo, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science*, 8, 101-105.
- Saffran, J. (2001). The Use of Predictive Dependencies in Language Learning. *Journal of Memory and Language*, 44, 493-515.
- Saffran, J. (2001b). Words in a sea of sounds: the output of infant statistical learning. *Cognition*, 81, 149-169.
- Saffran, J. (2002). Constraints on Statistical Language Learning. *Journal of Memory and Language*, 47, 172-196.
- Saffran, J. and Wilson, D. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, 4, 273-284.
- Spada, N. (1997). Form-Focused Instruction and Second Language Acquisition: A Review of Classroom and Laboratory Research. *Language Teaching*, 30, 73-87.
- Stadeler, M. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 3, 674-685.
- Stein M, Federspiel A, Koenig T, Wirth M, Strik W, Wiest R, Brandeis D, Dierks T. (2012). Structural Plasticity in the Language System Related to Increased Second Language Proficiency. *Cortex*, 48, 458-465.
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual Differences Reveal Stages of Language Acquisition. *Bilingualism: Language and Cognition*, 16, 367-382.
- Thompson, S., & Newport, E. (2007). Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning and Development*, 3, 1-42.
- Tincoff, R. & Jusczyk, P. W. (1999). Some Beginnings of Word Comprehension in Six-Month-olds. *Psychological Science*, 10, 172-175. Reprinted in D. Muir & A. Slater (Eds.), *Infant development: The essential readings* (pp. 270-278). Oxford, UK: Blackwell Publishers, 2000.
- Tokowicz, N & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Journal of Second Language Acquisition*, 27(2), 173-204.
- Toro, J., Sinnett, S., & Soto-Franco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97, B25-B34.
- Toro, J & Trobalon, J (2005). Statistical Computations over a Speech Stream in a Rodent. *Perception and Psychophysics*, 67, 867-875.
- Werker, J., & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual re-organization during the first year of life. *Infant Behavior and Development*, 7, 49-63.
- Wonnacott, E., Newport, E., & Tanenhaus, M. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, 56, 165-209.

6. Appendix

6.1 The lexicon

Nouns. There were ten two-syllable nouns [*flerbit*, *melnag*, *jernat*, *klamon*, *blifon*, *nagid*, *slagum*, *runpat*, *kowalt*, *tasnor*], each with a different stem-initial consonant and each conforming to English phonotactic constraints. There were no lexical restrictions on the noun's occurrence: each noun occurred freely with all verbs in all clause types.

Verbs. There were eight one-syllable verbs, split into intransitive (*flern*, *trag*, *glim*, *shen*) and intransitive (*loom*, *pemz*, *kofe*, *bund*) classes, with transitivity lexically specified and simply determined by the construction type that the verb appeared in. Verbal morphology consisted of an obligatory number-marking prefix (kV-, kVC-), which varied systematically with preceding nominal case markers (described in further detail below). The verbal system of the language also contained four auxiliary verbs (*iti*, *til*, *da*, *ada*) that appeared within the head-final verb phrase. Each of the eight main verbs obligatorily selected for one of the (lexically specified) four auxiliaries, such that there were two auxiliaries associated with the transitive class (*iti*, *ada*) and two associated with the intransitive class (*til*, *dal*).

Determiners. The language contained two determiners, a definite (*ba*) and an indefinite (*po*). There were no restrictions on the occurrence of determiners, all nouns could occur with either the definite or the indefinite.

Temporal adverbials. The language contained four temporal adverbials (*boke ton ol*: "yesterday," *boke ton ap*: "a long time ago," *fune ton ol*: "tomorrow," and *fune ton ap*: "in the distant future.") that always occurred sentence finally. There were no

restrictions on these adverbials: all occurred with all verbs in all constructions.

6.2 *The grammar*

Word Order and Basic Structures. The language had an SOV word order, and contained the two possible constructions in (1) and (2):

(1) [Noun Phrase 1] [Noun Phrase 2] [[AuxV-Transitive][V-Transitive]]^{VP} Adverbial

(2) [Noun Phrase 1] [AuxV-Intransitive][V-Intransitive]]^{VP} Adverbial

Construction (1) is a transitive clause, with [Noun Phrase 1] functioning as the subject and [Noun Phrase 2] as the object. Construction (2) is an intransitive clause with [Noun Phrase 1] the sole argument of the verb.

Case Marking and Number Agreement. The morpho-syntax of the artificial language consisted of a split ergative case marking/number agreement system. In an ergative language, subjects of intransitive verbs pattern with the direct objects of transitive verbs (and are marked with absolutive case), while subjects of transitive verbs are marked with ergative case (see Dixon, 1994 for a comprehensive review of ergativity). Split ergative languages are those in which some grammatical relations are characterized by this ergative/absolutive pattern, while others are in the more familiar nominative/accusative alignment that groups subjects together regardless of the thematic structure of the verb of which they are an argument. Split ergativity manifests in many ways cross-linguistically, but here the case system was in ergative/absolutive alignment and the number agreement system was in nominative/accusative alignment.

While this is a relatively rare pattern cross-linguistically, it is attested in many languages of the Austronesian language family (see Tauberschmidt & Bala, 1992 and Tauberschmidt, 1999 for a discussion of this pattern in Sinaugoro, a language of New

Guinea). The ergative/absolutive case system marked intransitive subjects and direct objects of transitive verbs with absolutive case (singular: *-il*, plural: *-amt*) and transitive subjects with ergative case (singular: *-ot*, plural: *-e*). The nominative/accusative number agreement system created a morpho-syntactic dependency between the main verb and the (either ergative or absolutive marked) subject. Number was marked on the verb through allomorphic variation of the prefix (kV-, kVC-). The vowel in this verbal prefix harmonizes with the unique vowel in the case-marking suffix on the agreeing noun,²⁰ and the consonant right-to-left reduplicated from the initial segment of the verb stem; with presence and absence of this consonant indicating singular and plural, respectively.

It is worth noting that, though both the agreement and subcategorization systems depended crucially on properties of the verb phrase, their instantiation at this position was entirely orthogonal. The agreement system was entirely dependent on the form of the number marking prefix, and the manner in which it co-varied with the case marker on the agreeing noun. The auxiliary and main verbs themselves were irrelevant to the functioning of the agreement system. In contrast, the subcategorization system, a system to which the form of the verbal prefix was entirely irrelevant, was entirely dependent on restrictions on the co-occurrence of auxiliary and main verbs, and main verbs and sentence types. It was this independence of the two grammatical systems, nonetheless encoded at the same critical positions, which permitted a clean manipulation of focal attention, one that allowed us to be sure that subjects were

²⁰ See, for example, the Somali language, for examples of vowel harmony across word boundaries. Though this pattern is rare, it is attested.

exposed to both grammatical systems while attending to only one.

CHAPTER 3: Elicitation of Early Negativity in Sentence-processing Contexts Depends on Attentional Efficiency: An Event-related Potential (ERP) Study

Abstract:

Here we examine the relationship between language and the brain's attention networks, testing the hypothesis that EN is not an underlying linguistic, nor purely sensory response, but rather reflects an attentional enhancement of the domain-general N100. Participants completed both a sentence-processing task, designed to elicit early negativity as well as other linguistic ERP components, and the Attention Network Task (ANT), which assesses the efficiency of participant's alerting, orienting, and executive attention networks. Efficiency-based median splits showed that early negativity was only elicited in participants with low efficiency orienting or low-efficiency executive networks, and a multiple regression analysis revealed that, in combination, the efficiency of these networks predicted a significant amount of variance in mean amplitude of early negativity. This pattern of results was in direct opposition to the patterns observed for other linguistic ERP components: all groups showed significant LAN, N400, and P600 effects that differed only in terms of effect magnitude and distribution at the scalp. Taken together, these results support our hypothesis, and suggest a domain-general interpretation of the functional significance of early negativity, in which it reflects an attentional modulation of the N100.

1. Introduction

1.1 Language and attention

A productive research strategy in the linguistic event-related brain potential (ERP) literature has been to assume that individual variation in electrical brain responses to language stimuli is a consequence of variation, not within the language system itself, but rather of individual variation in the domain-general cognitive systems that the language system interfaces with. In other words, it is variation in language “performance,” rather than language “competence,” with the latter argued to be largely invariant across speakers of a language (Chomsky, 1965). Linguistic ERP research operating under this assumption has thus far focused on the brain’s domain-general working memory system, a limited capacity system which functions to maintain and manipulate information over short time intervals (Baddeley, 1993). The most persuasive findings in this literature have shown that individual variation in working memory capacity, as assessed by the reading span task (Daneman & Carpenter, 1980), (sometimes) correlates with differences in ERP responses to linguistic stimuli (Munte et al., 1998; Vos et al., 2001; Gunter et al., 2003; Fiebach et al., 2004; Nieuwland & Van Berkum, 2006).¹ These observations tentatively suggest that the language system is tightly coupled with the brain’s working memory system(s).

While investigations into the language-working memory interface have been fruitful, recent research has begun to question the extent to which working memory should be considered an isolable neurocognitive system, and therefore whether working memory is an appropriate object of inquiry for such investigations. This line of

¹ It should be noted that the direction of correlations varies across studies.

research has raised the possibility that the relationship being mapped in the studies above is in fact a relationship between language and an emergent ‘cognitive ability,’ rather than an actual hard-wired neurocognitive system. Many researchers (Cowan, 1988; 1995; 1999; Ericsson & Kintsch, 1995; Garavan, 1998; Oberauer, 2002; Awh et al., 2004, Vasisth et al., 2005, 2006, among others) have recently argued that what has previously been described as working memory is better characterized as attention operating over the contents of long-term memory. Put simply, what has previously been described as the ‘contents of working memory’ and the capacity limits on these contents is better understood as ‘focusing attention on the contents of long-term memory,’² an operation underpinned by attentional pointers in prefrontal cortex and long-term memory representations in posterior temporo-parietal regions (Ruchkin, 2003). Even if one rejects the assertion that working memory is an emergent epiphenomenon (an admittedly strong claim), it has recently become clear that at a minimum the brain’s working memory and attention systems are tightly coupled (see Smyth & Scholey, 1994; Awh & Jonides, 1998; Awh & Jonides, 2001 (figure 4); Vogel et al., 2005 among others).

Here we hypothesize that an individual’s language processing efficiency is dependent on the efficiency of the brain’s attention networks. As outlined above, the relationship between attention and language is an appropriate object of inquiry because either (i) what has previously been characterized as the brains’ working memory system is actually an epiphenomenal ability resulting from the interactions of attention and

² Under this view, capacity constraints are located within the attention system, consistent with the data in Marois & Ivanoff (2005) that locates the “bottle-necks” of human cognition in nodes of the brain’s attention networks (see also Vogel & Machizawa, 2004; Vogel et al., 2005, and Luck & Vogel, 2013).

long-term memory, or (ii) the brain's working memory and attention systems are tightly coupled, and only the working memory aspect of these interdependent systems has been thoroughly investigated. Below we turn to a discussion of the early left anterior negativity (*eLAN*), an ERP component elicited in sentence processing contexts, with the goal of first justifying it as an appropriate dependent measure with which to investigate the language-attention interface, and subsequently developing an attention-based account of this brain response.

1.2 The eLAN

1.2.1 The traditional view

In order to investigate the relationship between language and attention, it is first necessary to select a brain response to linguistic stimuli that is putatively attentional in nature. This will allow us to determine whether the properties of this “linguistic” brain response co-vary with attentional efficiency, and therefore if what initially appears as a linguistic response is in fact attentional in nature. In this section we discuss such a response, the early left anterior negativity (*eLAN*), before developing an attention-based account of its functional significance that will motivate the current work.

Neville et al. (1991) were the first to report a negativity, maximal over the left hemisphere, with very early onset (peaking at 125 msec.), in response to the type of phrase structure violations in shown in (1b) below:

(1a) The scientist criticized Max's proof OF the theorem.

(1b) *The scientist criticized Max's OF proof the theorem.

Shortly thereafter, Friederici et al. 1993 reported a similar response in the auditory modality to comparable violations in German (literal English translation: **the friend was in the ___ visited*, “The friend was visited in the ___”), in this case with a pronounced left anterior distribution. In both of the manipulations described above, a word of the wrong grammatical category, a so-called ‘word-category violation’ (WCV) (in Neville et al. (1991) the preposition *of*; in Friederici et al. (1993) the verb *visited*) appears instead of the expected nominal head and therefore cannot be integrated into the parse under construction. On the basis of these observations (and numerous studies employing the same or similar experimental paradigm(s) (e.g. Weber-Fox & Neville, 1996; Friederici et al., 1999; Gunter et al., 1999; Hahne & Jeschniak, 2001; Kotz et al., 2003; and Hahne et al., 2004 among others) and the early latency of the effect, in 2002 Friederici dubbed this effect the early left anterior negativity and argued that it indexed an initial informationally encapsulated “first pass” stage of a modular and serial parser, during which initial syntactic structure is built on the basis of word category information alone (Frazier & Fodor, 1978).

However, alongside this research an accumulation of observations began to emerge suggesting that the functional interpretation of the eLAN proposed in Friederici (2002) was incorrect. Subsequent to the original studies, it has been noted that eLAN-like effects can be elicited by unexpected articles in *fully grammatical* gapping constructions (Kaan et al., 2004), by lexical items that are fully grammatical but pragmatically or collocationally odd (Moreno et al., 2002; Rosenfelt et al., 2009), and by words (Roll et al., 2007) or bound morphemes (Zhang & Zhang, 2008) that are ungrammatical but nonetheless of the *correct* word/morphological category. Additional

studies have shown that WCVs that swap nouns for verbs and vice versa (Federmeier et al., 2000; Rosenfelt et al., 2011) elicit an N400-P600 complex rather than *eLAN*. In sum, these studies demonstrate that WCVs are neither necessary nor sufficient for the elicitation of *eLAN*, and therefore that a novel functional interpretation with broader empirical coverage is required.

The sensory hypothesis of the functional significance of *eLAN* was first proposed in Dikker et al. (2009) and further refined in Dikker, 2010 and Dikker & Pylkaanen, 2011). This hypothesis was based on the results of a magnetoencephalography (MEG) study that employed the same violations in Neville et al., 1991, and Friederici et al. 1993, and localized M100 effects (the magnetic equivalent of the domain-general N100, an effect to which we return shortly) to primary visual cortex. Based on the localization of these effects, the authors reinterpreted the *eLAN* as a domain-general sensory response to low-level physical features of the linguistic input, features that they argue are actively predicted during sentence processing. The authors argue that this response is ontologically sensory, rather than a syntactic or even linguistic. Under this account, early negativities index sensory matching operations computed on low-level physical features:³ when features predicted don't match features encountered, early negativity is elicited. Note that this account can both (i) explain the early latency of the effect as it reflects processes occurring early in

³ The nature of the linguistic level(s) at which these form-based predictions occur is still under investigation, but at a minimum, predictions seem to be generated at the level of category-signaling closed class morphology (Dikker et al., 2009), orthographic and/or phonological features typical of grammatical categories (Dikker et al., 2010), and lexico- semantic features (Dikker & Pylkaanen, 2011).

the visual processing stream, and (ii) provide broad empirical coverage, accounting for the entirety of the results discussed above.

1.2.2 Early negativities in sentence-processing contexts: An emerging view

Largely because of the influence of the work of Dikker and colleagues, claims about the functional significance of early negativities elicited in sentence processing contexts are clearly in need of revision. Because of its intuitive appeal and broad empirical coverage, the sensory hypothesis has generated a great deal of research dedicated to mapping the antecedent conditions under which these effects are reliably elicited, a productive and important research enterprise. While the research discussed above is consistent with the hypothesis proposed by Dikker and colleagues, the current study can perhaps further articulate the nature of the sensory matching processes at its core. We feel that the sensory hypothesis is lacking an explanation of the neurocognitive mechanism by which form-based predictions are generated and matched against the input in sensory cortex.⁴ Additionally it lacks an explanation for why a form-based mismatch between observed and predicted input would manifest as enhanced negativity over sensory cortex. As such, the current study is designed to articulate further the processes first proposed in the sensory hypothesis.

We hypothesize that the effect previously referred to in the neuro-linguistic literature as the *eLAN* is actually an attentional modulation of the domain-general N100,

⁴ Dikker et al. (2009, 2010) discuss two alternative mechanisms via which form-based predictions could be generated and matched against the physical features of linguistic input: either (sub-)lexical representations are stored in sensory cortex allowing predictions to be both generated and matched against the input *internal* to sensory cortex, or predictions are generated *externally* and fed to sensory cortex in a top-down fashion enabling sensory cortex to perform the matching operation. Even in their most recent work the authors do not distinguish between these alternatives.

the earliest sensory response that is modulated by manipulations of selective attention (see Hillyard et al., 1973), and as such we will no longer refer to these responses as eLAN effects. It is our hypothesis that this attention-based account of early negativities elicited by WCVs, generated from a review of the literatures on the brain's attention networks and ERP studies of selective attention, can address the areas in which the sensory hypothesis lacks explanatory power as described above. Below I provide a brief review of these literatures in order to develop a fully articulated attention-based account of the functional significance of early negativities elicited in sentence-processing contexts.

1.3 The attention networks of the human brain

Research investigating the human attention system, which functions to select sensory information rapidly for enhanced processing, dates back to at least the 1890s, and thus a comprehensive review of the topic is beyond the scope of this paper. For present purposes, we rely on an influential description of the three sub-systems of the brain's attention system as described in Posner & Peterson (1990). The task used in the current study (the Attention Network Task or ANT; Fan et al., 2002) is designed to assess the efficiency of these three attentional sub-systems.

The alerting system: The first step in identifying behaviorally relevant stimuli for enhanced processing is maintaining an alert and vigilant state that enables the subsequent selection of the stimulus of interest. The maintenance of this alert state is critical for subsequent orientation of the attention system to these stimuli.

The orienting system: Attentional orienting is the process by which a stimulus is brought into the center of attention for selective processing, and can occur either overtly

(i.e. accompanied by visual saccades) or covertly (via a shift of the so-called “attentional spotlight”). Orienting to a spatial location produces more rapid behavioral responses (Posner et al., 1980) and larger ERP responses to stimuli that have been oriented to (Mangun & Hillyard, 1987), as well as enabling the detection of targets at lower contrast thresholds in visual search paradigms (Downing, 1988). Generally speaking, orienting increases processing efficiency. This attentional sub-system satisfies the functional requirement of responding to behaviorally relevant stimuli that were previously outside the focus of attention by bringing them into attentional focus. Behavioral goals can trigger orienting responses when behaviorally relevant targets appear, stimuli that would otherwise be unimportant and non-distinctive (a behavior that Corbetta & Shulman (2002) refer to as “contingent orienting”). Across a variety of paradigms, activation of this network is not driven by stimulus salience, intensity, or distinctiveness, but rather by behavioral relevance (Corbetta et al., 2000; Braver et al., 2001; Downar et al., 2001, Bledowski et al., 2004; Indovina & Malacuso, 2007).⁵ Corbetta & Shulman (2002) propose that the function of this network is to function as a “circuit breaker” that interrupts ongoing cognitive operations and shifts attention to an unexpected, but behaviorally relevant, stimulus, a function that is crucial to our attention-based account of early negativity outlined below.

The executive attention system: The subsystem of the brain’s attention network that is under cognitive control and biases the selection of incoming input has variously

⁵ Note that the oddball paradigms employed in some of these studies do not involve a spatial component, providing further support for the notion that the orienting network does not control only (re-)orienting of a spatial nature (Kirino et al., 2000; see also Marois et al., 2000). This is crucial for the account developed here, as in most sentence processing experiments all stimuli are presented centrally.

been described as controlling “top-down attention,” “goal-directed attention,” or “conscious attention.” In the cueing paradigms that are often used to investigate the functioning of this attentional system, participants are first presented with a cue that specifies the location of an upcoming target, followed by a delay period during which no visual stimulus is present, followed by target presentation. During the delay period, subjects are controlling attention in a top-down fashion in order to anticipate the appearance of the targets, either at cued locations or with specific features. This system is responsible for generating signals based on the current set of behavioral goals (i.e. detect, and respond to, the target), signals that bias the processing of goal-related stimulus features or locations (Kastner et al., 1999; Shulman et al., 1999; Corbetta et al., 2000; Hopfinger et al., 2000; Serences et al., 2001; Serences et al., 2005; Silver et al., 2007). As such, the executive attention system functions to modulate activity in sensory areas in advance of a predicted target. Acting in concert with the orienting system, the executive system functions to generate predictions about the input, modulate sensory areas in anticipation of the target stimulus, and then subsequently trigger orienting responses to unexpected but behaviorally relevant input.

It appears to be the case that the actions of the executive and orienting attention networks can further elucidate the nature of the sensory matching operations described in the sensory hypothesis, i.e. the interactions between these systems can account for the manner in which form-based predictions are generated by the executive attention system (predictions that subsequently modulate sensory responses) in sentence-processing contexts, subsequently triggering orienting responses to stimuli possessing behaviorally relevant features. In other words, it is the interaction of these attentional

processes that we hypothesize leads to larger N100 effects in response to WCVs.

Next we discuss ERP studies of selective attention in order to provide an account of why a mismatch between features that are predicted and those that are encountered would elicit enhanced negativity over sensory cortex. Lastly, we propose a fully developed attentional account of N100 effects elicited in sentence processing contexts.

1.4 ERP studies of selective attention

The ERP literature on selective attention is vast, dating back to the early work of Steven Hillyard in the early 1970s (see Näätänen, 1979, 1982 for a review of this early work, and Hopfinger et al., 2004 for a review of more recent studies). Therefore, the goal of this section is not to review this literature comprehensively, but rather to discuss briefly a handful of studies that demonstrate, across varying experimental paradigms, that selectively attended stimuli elicit larger N100 responses compared to their unattended (but physically identical) counterparts.

In the first ERP experiment designed to investigate the effects of selective attention on stimulus processing, Hillyard et al. (1973) had subjects perform a dichotic listening task in which they were instructed to attend to tones in one ear (while counting instances of one tone type) while ignoring tones played to the other ear. The authors reported enhanced negativity onsetting around 60-70 msec. (the N100 effect, one of the first exogenous ERP components modulated by selective attention) in response to attended tones when compared to their physically identical but unattended counterparts, demonstrating that selective attention to a stimulus leads to larger N100 responses. This early work led to an explosion of research (Eimer 1996; Anllo-Vento & Hillyard 1996;

Luck et al. 1993; Vogel & Luck 2000; Martinez et al., 2001, among many others) investigating the effects of selective attention on early brain responses, with subjects typically instructed to attend to a location in space (while ignoring other locations) or to a particular stimulus feature (while ignoring others). Most pertinent to the current work, Luck et al. (1993), Eimer (1996), Anllo-Vento & Hillyard (1996), and Vogel & Luck (2000), had subjects make target discriminations based on physical features. Larger N100 responses were elicited in all of these studies when stimuli with to-be attended features were compared to those without. As such, it appears that regardless of the nature of the paradigm employed, selective attention to a stimulus results in a larger N100 response.

1.5 An attention-based account of early negativities elicited in sentence processing contexts

Based on the preceding reviews of the literature on the brain's attentional networks and ERP studies of selective attention, it is now possible to develop a full attentional account of the early negativities elicited by WCVs with the goal of adding to the explanatory power of the sensory hypothesis of Dikker and colleagues.

During sentence processing experiments, participants are required either to make grammaticality judgments or to answer comprehension questions, and as such their behavioral goal is to process the sentence in order to successfully perform these tasks.

The alerting network maintains the alert state necessary for task performance.

Concurrently, the executive attention network is generating predictions about physical features of the incoming linguistic input (see, for example, DeLong et al., 2005) and continuously modulating activity in sensory cortex in an anticipatory fashion (e.g.

Serences et al., 2005; Silver et al., 2007). Based on the behavioral goals of the participant, and the expectations generated by the executive attention network, the orienting network orients to unexpected but task-relevant features of the input, in this case word category violations (WCVs). This orienting response leads to selective processing due to the deployment of additional attentional resources, and N100 amplitude increases. It is this increase in amplitude of the N100 response that has been referred to in the neuro-linguistic literature as *e*LAN. The present study was designed to evaluate this hypothesis.

2. Methods

2.1 Participants

40 monolingual native speakers of English participated in the experiment (24 females, 16 males), completing both a sentence processing task and the attention task described below (see section 2.1.1). Participants received either course credit or \$9 an hour for their participation. All participants were between the ages of 18-24 (mean 19.2), and were enrolled as students at the University of California, San Diego. All participants were right-handed with no neurological or psychiatric disorders and normal or corrected-to-normal vision.

2.2 Materials: Sentence processing experiment

We adopted the materials for the sentence processing experiment from previous studies that reliably elicited *e*LAN, LAN, N400, and P600 effects (see TABLE 3.1). The comparison between WCVs (condition A) and their controls (condition B) is known to elicit early negativity between 50-150 msec. (Neville et al., 1991, Friederici et al., 1993); the comparison between object relative clauses (condition C) and their

coordinate clause controls (condition D) is known to elicit LAN between 300-500 msec. at the main clause verb (King & Kutas 1995, among others); the comparison between semantic anomalies (condition E) and their controls (condition F) elicits N400 effects between 300-500 msec. (Kutas & Hillyard, 1983; Neville et al., 1991); and the comparison between WCVs (condition A) and their controls (condition B) reliably elicits P600 effects between 500-800 msec. in addition to early negativity (Neville et al., 1991). As originally constructed the materials also contained a comparison between transitive constructions and their intransitive controls (see Osterhout & Holcomb, 1992), but this comparison elicited no P600 effects in our group of participants and as a result will no longer be discussed. Participants read 40 sentences of each of the 8 experimental conditions, and the 320 experimental sentences were pseudo-randomized into 2 lists. More than two sentences of the same condition never appeared consecutively and there were never more than five items from the same condition within a sequence of ten items.

TABLE 3.1: Experimental sentences: Sentence-processing experiment

Condition	
A. *WCV	The scientist criticized Max's <i>of</i> proof the theorem.
B. Control	The scientist criticized Max's proof <i>of</i> the theorem.
C. Object Relative	The soldier who the sailor roughly pushed <i>smashed</i> a bottle against the bar.
D. Control	The soldier roughly pushed the sailor and <i>smashed</i> a bottle against the bar.
E. Semantic anomaly	The judge read Tom's <i>stories</i> about crime.
F. Control	The judge read Tom's <i>tones</i> about crime.

2.2.1 Attention Network Task (ANT)

The ANT (originally appearing in Fan et al., 2002) was designed to assess the efficiency of an individual's alerting, orienting, and executive attention networks (see section 1.3) independently of each other. A trial of the task is schematized in FIGURE 3.1 below:

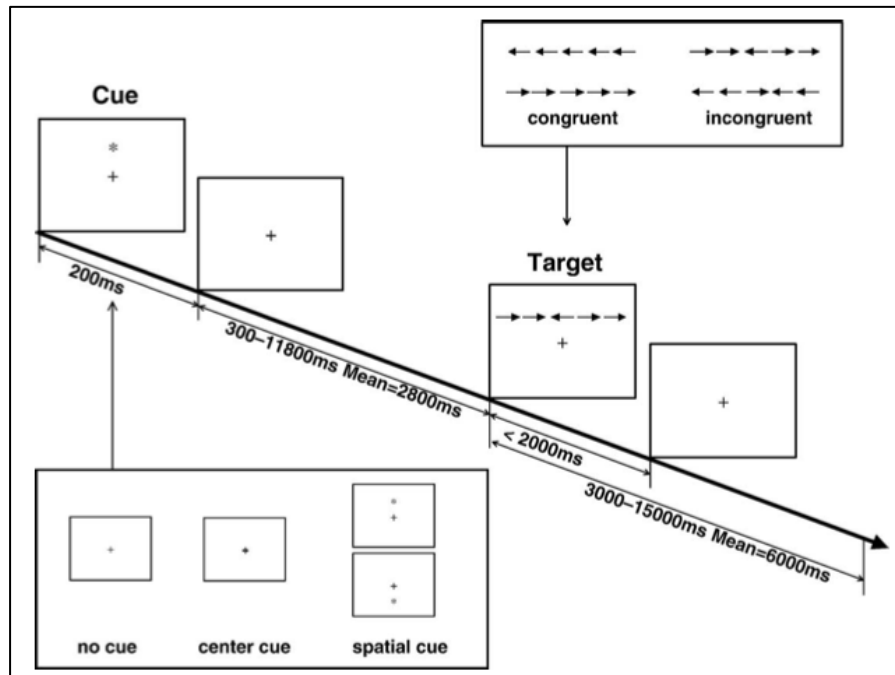


FIGURE 3.1: Schematic of the Attention Network Task (ANT)

During a trial of the ANT, the participants' task is to respond as rapidly as possible to congruent (<<<< or >>>>) or incongruent (<<>< or >><>) flanker targets (see Eriksen & Eriksen (1974) for a description of a simple Flanker task). Prior to target presentation, subjects view either a simple fixation cross that provides no information about the upcoming target, a central cue occurring at fixation (*) that alerts the subject to the upcoming presentation of a target (but not its location), or a spatial cue (*) that

alerts the subject to the upcoming presentation of the target as well as its spatial location (either above or below fixation; see Figure 3.1).

2.3 Procedures

The sentence processing experiment, which participants completed first, consisted of five blocks, each containing 64 sentences and lasting approximately 15-17 minutes. Participants were run in a single EEG session that lasted approximately 2.5 hours, including preparation. During the session, subjects were seated comfortably in a chair in a sound-attenuated booth. 1500 msec. before the first word of a sentence, a red square appeared in the middle of the screen and remained throughout sentence presentation for fixation purposes. Sentences were visually presented above fixation, with each word presented for 300 msec. (500 msec. stimulus onset asynchrony (SOA)). 2000 msec. after the end of the sentence subjects made grammaticality judgments, Response hands were balanced across participants to control for handedness. The next stimulus sentence began 3000 msec. after the participant's response. In order to familiarize the participants with the task, they completed a ten-sentence practice session before beginning the experiment.

The ANT consisted of 288 trials, which were broken into three blocks of 96 trials. The 288 total trials consisted of 96 trials for each cue type (no cue, alerting cue, and spatial cue), with each cue occurring equally often with congruent and incongruent targets and trial types balanced within blocks. Participants were instructed to respond as quickly and accurately as possible by indicating the direction in which the central arrow of the flanker target was pointing. In order to be consistent with the existing literature, we adopted the visual display parameters used in Neuhaus et al. (2009, 2010). The

central fixation cross (0.37 of visual angle subtended) was present throughout the experiment. Alerting cues (*) appeared at fixation, while spatial cues (*) appeared at 1.01 degrees above or below fixation and always accurately predicted the location of the upcoming target. Target stimuli consisted of five horizontally arranged arrows (3.28 degrees of horizontal visual angle) were presented at 1.01 degrees above or below the fixation cross (i.e. at the same position within the visual field as the spatial cues occurred).

2.4 Electrophysiological recording

The electroencephalogram (EEG) was recorded from 26 tin electrodes mounted geodesically in a commercially available Electro-cap. These sites included midline prefrontal (MiPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right medial frontal (LMFr and RMFr), left and right dorsal frontal (LDFr and RDFr), left and right medial central (LMCe and RMCe), midline central (MiCe), left and right medial dorsal central (LDCe and RDCe), left and right lateral temporal (LLTe and RLTe), left and right dorsal lateral parietal (LDPa and RDPa), midline parietal (MiPa), left and right lateral occipital (LLOc and RLOc), left and right medial occipital (LMOc and RMOc), and midline occipital (MiOc). Each electrode was referenced online to the reference electrode at the left mastoid and later re-referenced offline to the average of the two mastoids. To monitor blinks and eye movements, electrodes were placed on the outer canthi and under each eye. Impedances were kept below $5K\Omega$ during recording. The EEG was amplified using James Long amplifiers with an online band-pass filter (.01 to 100 Hz), and digitized at a sampling rate of 250 Hz.

2.5 Data analysis

2.5.1 Sentence-processing task

For the analysis of brain responses elicited during the sentence processing experiment, mean amplitude measurements were taken of single-word averages, which consisted of 1000 msec. epochs, including a 100 msec. pre-stimulus baseline. The time windows in which these measurements were taken were as follows: 50-150 msec. (N100 time window), 300-500 msec. (LAN and N400 time window), and 500-800 msec. (P600 time window). Trials contaminated by excessive muscle activity, amplifier blocking, or eye movements were discarded before averaging. This resulted in 8.9% of trials being rejected across conditions. The averaged data were algebraically re-referenced to the average of the activity at the two mastoid sites. ERP waves were smoothed offline using a low-pass filter with a 7 Hz cutoff for visualization purposes. Because of the design of the materials, a full factorial ANOVA was deemed inappropriate.⁶ Rather, we conducted four separate one-way ANOVAs to make the critical comparisons between (i) WCVs (condition A) and their grammatical controls (condition B) (EN and P600), (ii) object relatives (condition C) and their coordinate clause controls (condition D) (LAN), and (iii) semantic violations (condition E) and their non-anomalous controls (condition F) (N400). Each analysis had the repeated measures of experimental condition (2 levels) and electrode (26 levels). This will be referred to as the full analysis. In addition to the full analysis, a distributional analysis

⁶ This type of analysis would force us to compare across violation types and would lead to many false positive significant effects based on multiple confounds related to differences in lexical class and sentence position.

was conducted for each of the four main comparisons, including experimental condition (2 levels), hemisphere (2 levels: left and right), laterality (2 levels: lateral and medial) and anteriority (4 levels: prefrontal, frontal, temporal, and occipital) as factors.

Electrodes included in the distributional analysis were left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right lateral temporal (LLTe and RLTe), left and right medial dorsal parietal (LDPa and RDPa), left and right lateral occipital (LLOc and RLOc), and left and right medial occipital (LMOc and RMOc). The distributional analysis is referred to when it was necessary to resolve condition x electrode interactions in the full analysis. Furthermore, when it was necessary to corroborate small local effects, ANOVAs were performed on individual levels of these factors.⁷ Analysis of individual regions was only conducted when motivated by the presence of significant (or, at a minimum, marginal) interactions in the distributional analysis. The Huynh-Feldt (1976) correction for lack of sphericity was applied, and corrected *p*-values are reported with the original degrees of freedom.

2.5.2 ANT data

The efficiency of the alerting, orienting, and executive attention networks assessed by the ANT was calculated using the following subtractions. All effects were assessed for the participant group as a whole as well as for individual participants.

Alerting effect: RT (no-cue) *minus* RT (center-cue). The logic of this subtraction is that the center cue alerts the participant that the target is about to appear,

⁷ We did not conduct an ANOVA with a group factor in order to compare across efficiency groups due to established issues with comparing across different subject groups.

but gives no information as to the spatial location of the target (thus preventing the top-down control of attention leading to a (spatial) orienting response). Therefore, RT differences between these conditions, henceforth referred to as “alerting effects,” index the efficiency of the participant’s alerting subsystem: the larger the difference, the more efficient the network.⁸

Orienting effect: RT (center-cue) *minus* RT (spatial-cue). The logic of this subtraction is that, in both conditions, the subject is cued, but only in the spatial cue condition can the subject orient to the location in space at which the target will appear, thus reducing reaction time. Thus, RT differences between these conditions, henceforth referred to as “orienting effects,” are indicative of the efficiency of the participant’s orienting system, following the same logic described in fn. 8.

Executive attention effect: RT (incongruent) *minus* RT (congruent). The logic of this subtraction is that it is easier to respond to the central arrow when the flanking arrows are pointing in the same direction (congruent) than when they are pointing in the opposite direction (incongruent). RT differences in this comparison, henceforth referred to as “executive attention effects,” are thus indicative of the efficiency of the subject’s executive attention network, following the same logic described in fn. 8.

⁸ For example, if subject #1 takes 100 msec. to respond to the target when no cue is present, and 90 msec. to respond after the alerting cue, this subtraction will yield a value of 10 msec. If subject #2 takes 100 msec. to respond to the target in the no-cue condition, but only 50 msec. to respond after the alerting cue, this subtraction will yield a value of 50 msec. This larger difference for subject #2 shows that this subject “benefited” more from the alerting cue, and therefore that the alerting network of subject #2 is more efficient than that of subject #1. This basic logic holds for all ANT subtractions.

2.5.3 *Statistical analysis*

We first obtained measures of the efficiency of each of the three attention networks by conducting the subtractions described above, yielding individual alerting, orienting, and executive attention effects for each individual participant. Based on the size of these differences, we conducted an RT-based median split that divided the participants into 6 groups: 2 groups (high and low) x 3 attention networks (alerting, orienting, and executive), with 20 participants per group. We then analyzed each of the brain responses elicited during the sentence processing experiment to determine the extent to which these responses differed across high and low efficiency groups. Secondly, we conducted exploratory correlation- and regression-based analyses, attempting to determine the extent to which (i) participants' attentional efficiency correlated with the amplitude of EN at a subset of contiguous fronto-central electrodes (MiPf, LLPf, RLPf, LMPf, RMPf, LDFr, RDFr, LMFr, RMFr, LLFr, RLFr) that we selected based on visual inspection of individual participants' data,⁹ and (ii) we could construct a multiple regression model, including alerting, orienting, and executive attention effects (and subsets thereof) as predictors of EN mean amplitude at the electrodes above. While we acknowledge the exploratory nature of these analyses, we felt that they had the potential to provide additional evidence in support of our attentional efficiency-based hypothesis of EN.

⁹ We are aware of no precedent in the literature for an analysis of this type, and were hesitant to select different sets of electrodes for each individual participant based on where the response to WCVs was maximal, as we felt this constituted bad statistical practice. In lieu of this approach, we simply selected a consistent and contiguous set of electrodes across subjects. Proper statistical procedures for this type of analysis remain a topic for future research.

2.5.4 Predictions

If our hypothesis as to the functional identity of early negativities elicited in sentence-processing contexts is correct, we predicted that high and low attentional-efficiency groups as assessed by the ANT (in terms of either alerting, orienting or executive efficiency) would show differences in terms of the physical parameters of the N100 response. Crucially, if our hypothesis is correct, any observed differences in the parameters of the early negativities elicited by WCVs must be non-identical to the types of modulations of other ERP components elicited. In other words, if *all* “linguistic” brain responses vary across groups in a similar fashion, we would be unable to make any strong claims that what have previously been referred to as *e*LAN effects are in fact reducible to attentional modulations of the domain-general N100.

For the ANT data, we expected to replicate previous findings in studies using this paradigm: variability in attentional efficiency within and across subjects, decreasing RT with increasing cue informativity (RT: no cue > alerting cue > spatial cue), and decreases in RT when processing congruent compared to incongruent targets. For our correlational and regression-based analyses, if the predicted patterns in the median split analysis did in fact obtain, we predicted that we would see supporting evidence in the form of correlations between attentional measures and the parameters of early negativities elicited by WCVs, and that the results of our multiple regression model(s) would show evidence that some aspect(s) of attentional efficiency significantly predicted the mean amplitude of early negativity.

3. Results

3.1 ANT results

A summary of the results of the ANT is given in FIGURE 3.2 below.

	No Cue	Alerting Cue	Orienting Cue	Mean
Congruent				
Reaction Time	608.5 (126)	574.9 (128)	498.6 (158)	560.7 (126)
Incongruent				
Reaction Time	712.1 (142)	679.7 (19)	630.6 (119)	674.1 (158)

FIGURE 3.2: Mean RT for ANT task performance (SD in parentheses)

Overall, we replicated findings found in previous studies using the ANT. Results of two-tailed t-tests performed on the entire participant group's RT data showed that alerting ($t(1,39) = 3.6$, $p < .0001$), orienting ($t(1,39) = 9.02$, $p < .001$), and executive effects ($t(1,39) = 19.53$, $p < .001$) were all highly significant and in the expected direction. In other words, the data showed RT declines with increasing cue informativity (horizontal rows in Table 3.2 and x-axis in Figure 3.2), and independent of cue type a reduction in RT to congruent targets compared to incongruent ones (vertical columns in FIGURE 3.2 and the difference between the red and blue lines in FIGURE 3.2).

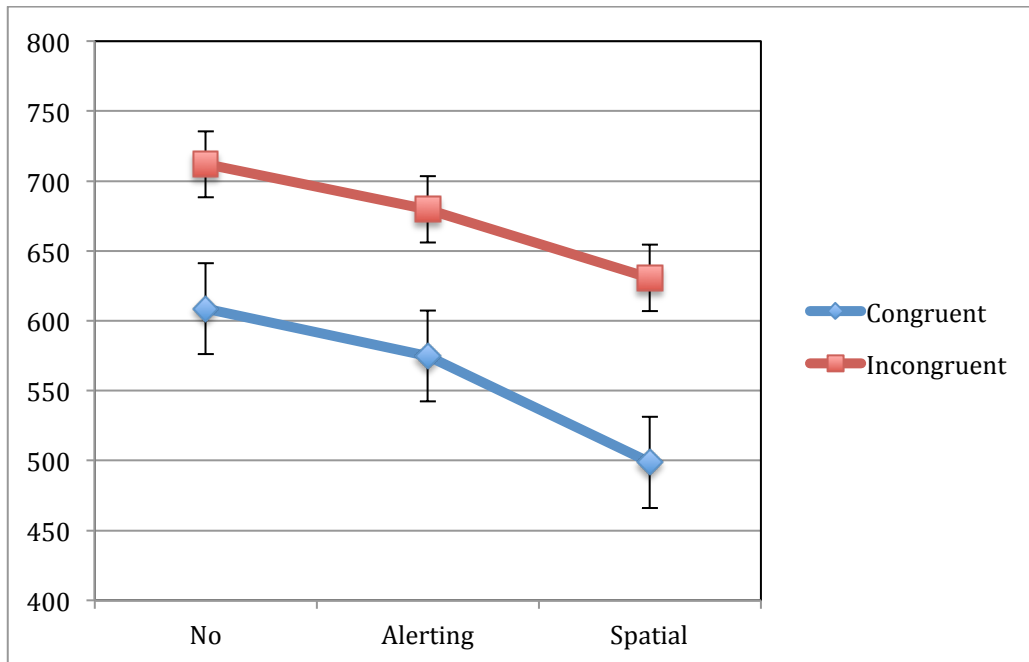


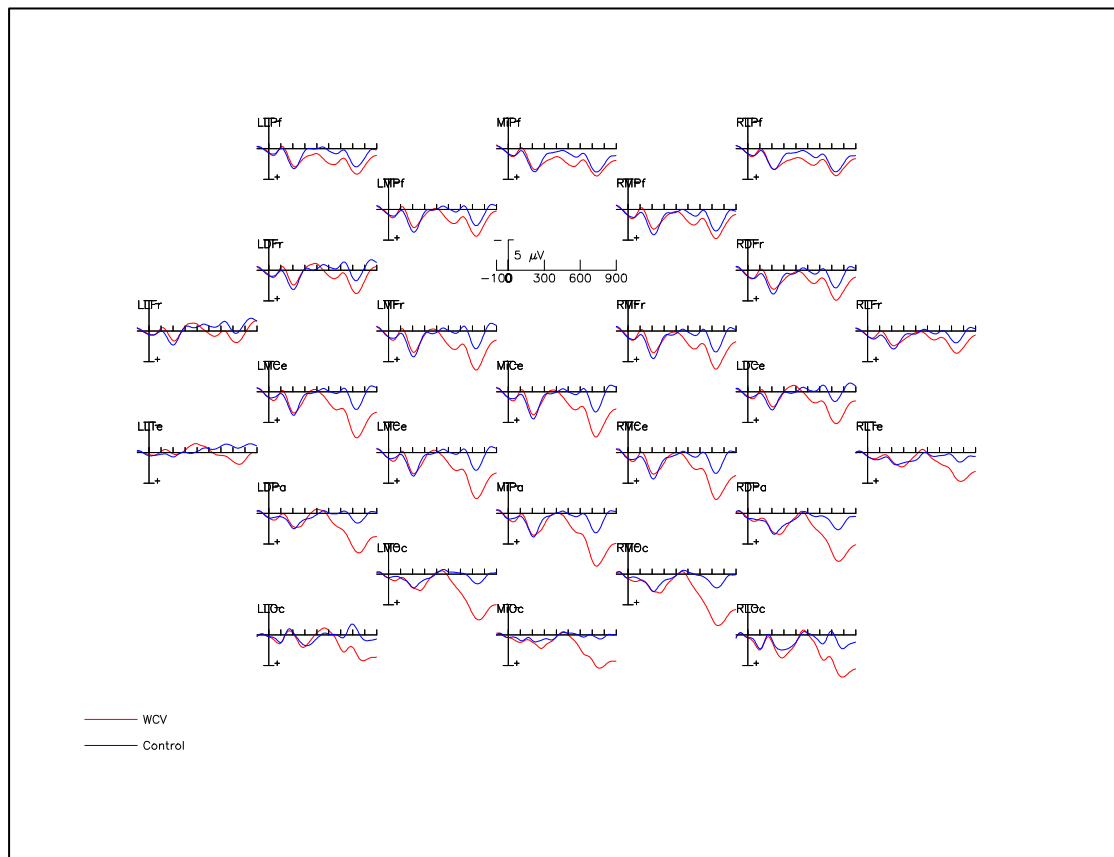
FIGURE 3.3: Plot of mean RT for ANT performance

We subsequently conducted median splits on these data based on the size of participants' alerting, orienting, and executive efficiency effects, yielding six groups of twenty individual participants. We conducted one-tailed t-tests that showed significant differences between the high (mean = 137.6 msec., S.D. = 55.9 msec.) and low (mean = 57.2 msec., S.D. = 37.4 msec.) efficiency alerting groups ($t(1,39) = 5.39$, $p < .001$), between the high (mean = 70.9 msec., S.D. = 38.9 msec.) and low (mean = -4.9 msec., S.D. = 35.4 msec.) efficiency orienting groups ($t(1,39) = 56.9$, $p < .001$), and between the high (mean = 172.6 msec., S.D. = 67 msec.) and low (mean = 56.8 msec., S.D. = 39.3 msec.) efficiency executive groups ($t(1,39) = 5.82$, $p < .001$). These results suggested a large degree of separation between the high and low efficiency groups.

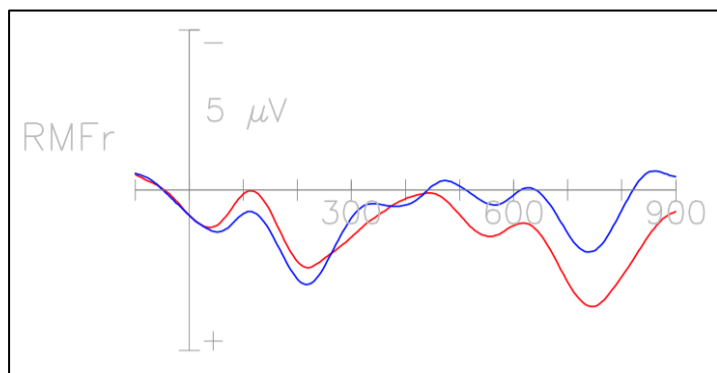
3.2. Sentence processing experiment:

The results of the omnibus and distributional ANOVAs are summarized in TABLES 3.3 and 3.4, respectively (see Supplementary Materials). The results of the omnibus ANOVA, conducted on all participant groups, showed robust main effects as well as condition x electrode interactions for LAN, N400, and P600 effects, as well as a significant condition x electrode interaction in the N100 time window (see FIGURES 3.4 – 3.7 below).

(A)



(B)



(C)

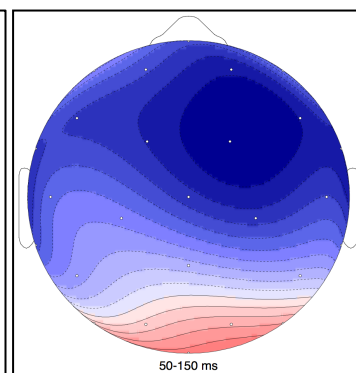
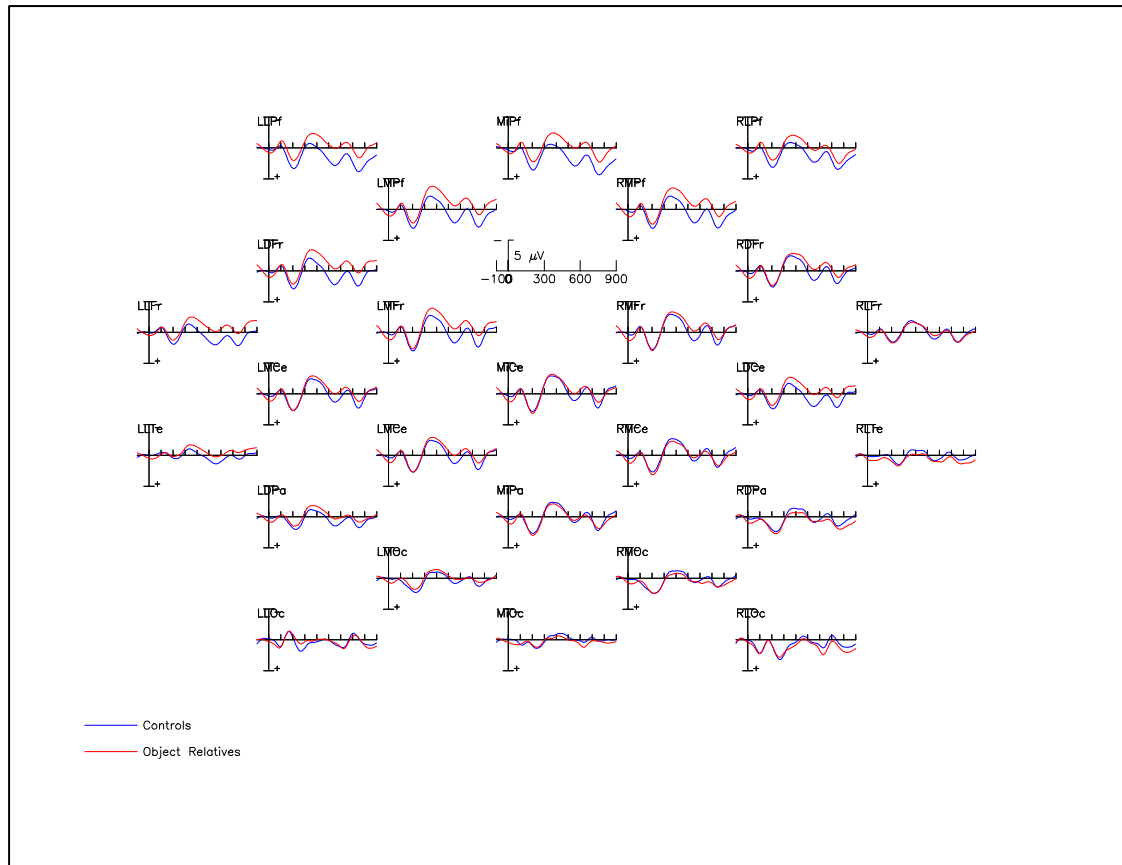
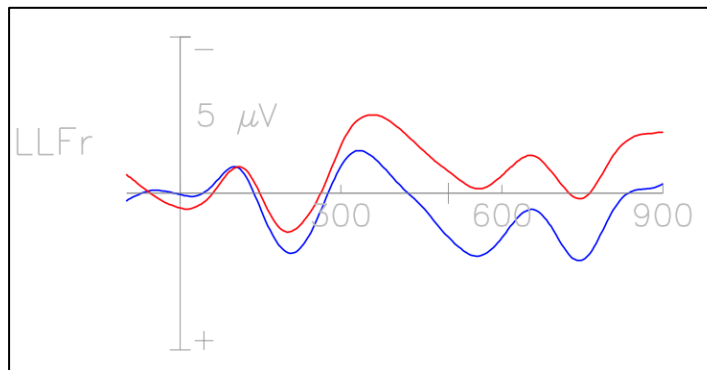


FIGURE 3.4: All participants. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode RMFr (0-500 msec.). (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec. Blue shading indicates enhanced negativity in response to WCVs.

(A)



(B)



(C)

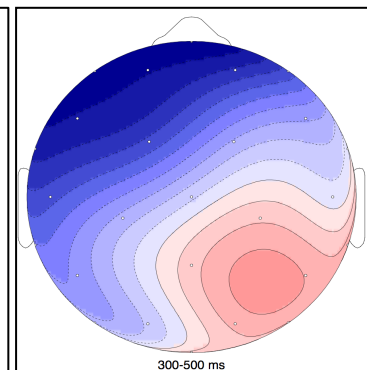
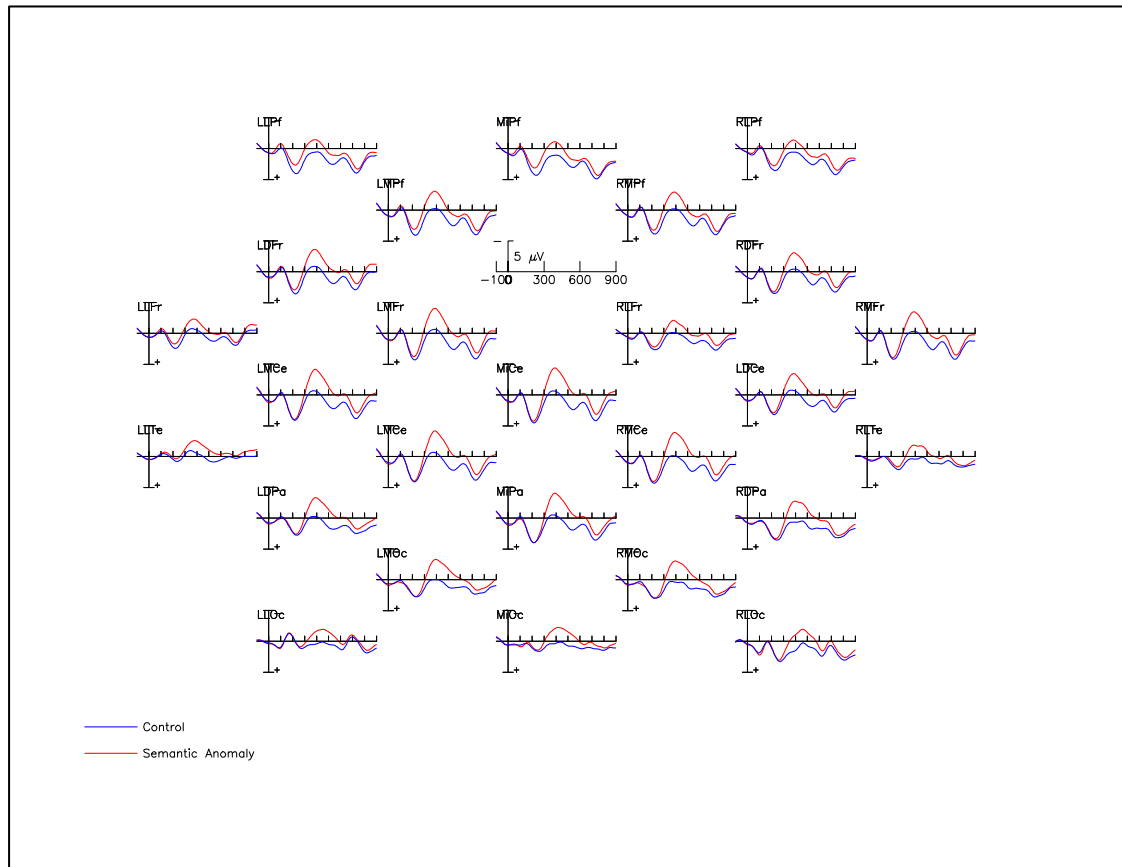
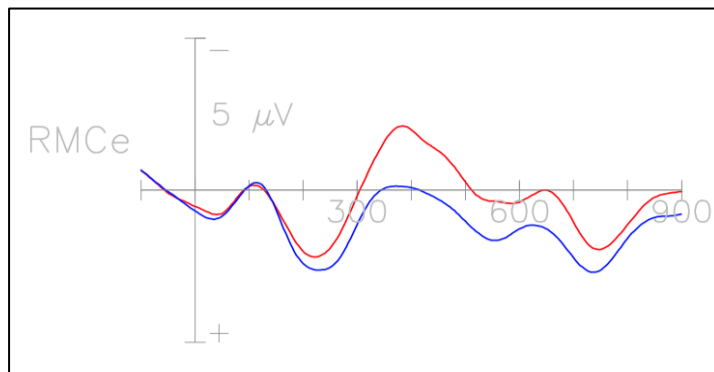


FIGURE 3.5: All participants. (A) Grand average ERP waveforms for **Object Relatives** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LLFr, (C) Topographic isovoltage map of the difference between conditions between 300 and 500 msec. Blue shading indicates enhanced negativity in response to object relative clauses.

(A)



(B)



(C)

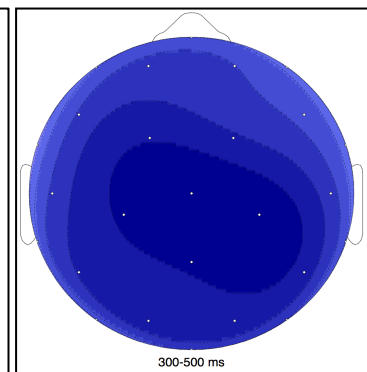
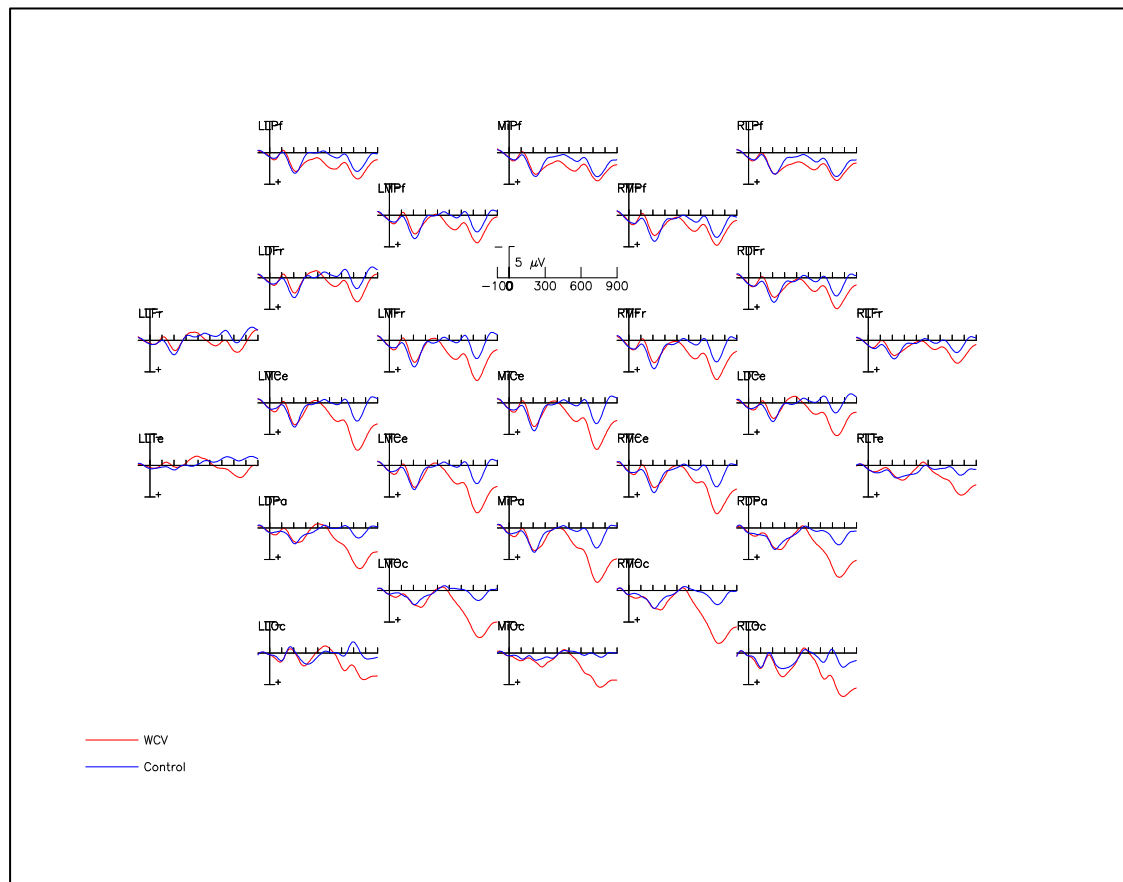
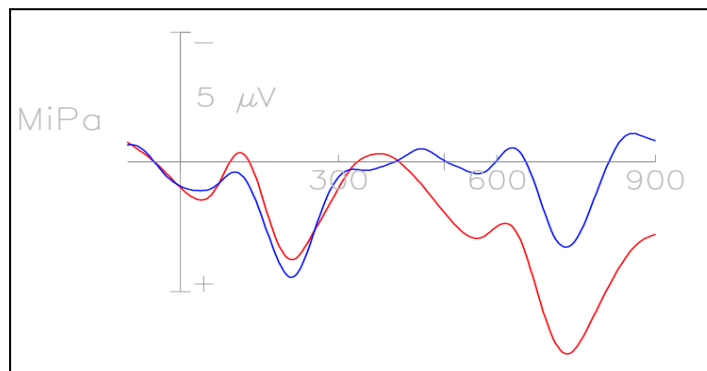


FIGURE 3.6: All participants. (A) Grand average ERP waveforms for **Semantic Anomalies** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LLFr., (C) Topographic isovoltage map of the difference between conditions between 300 and 500 msec. Blue shading indicates enhanced negativity in response to semantic anomalies.

(A)



(B)



(C)

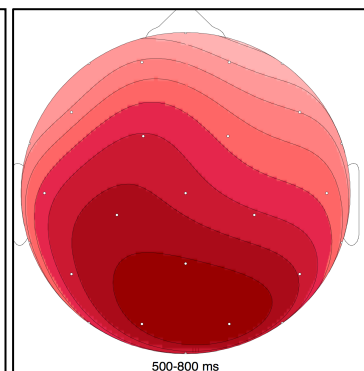
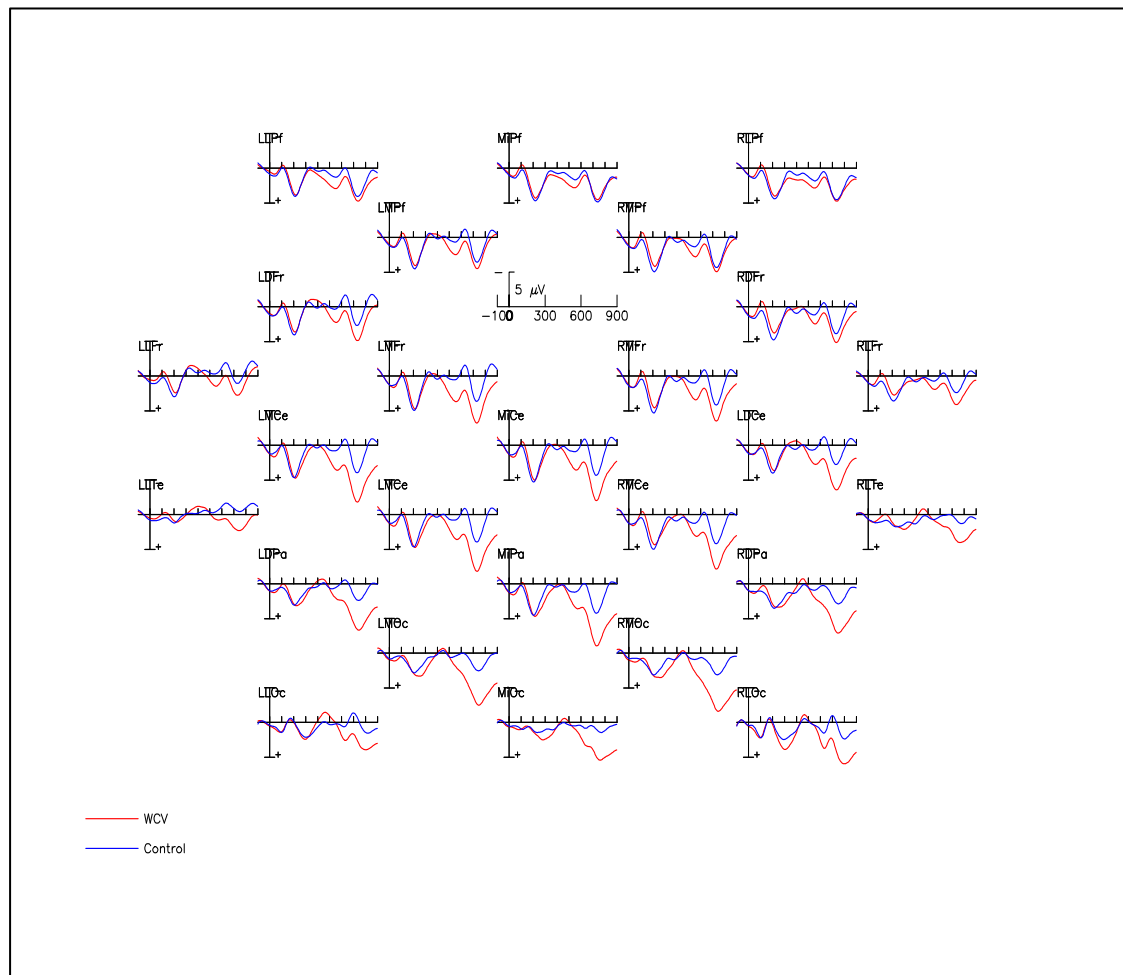


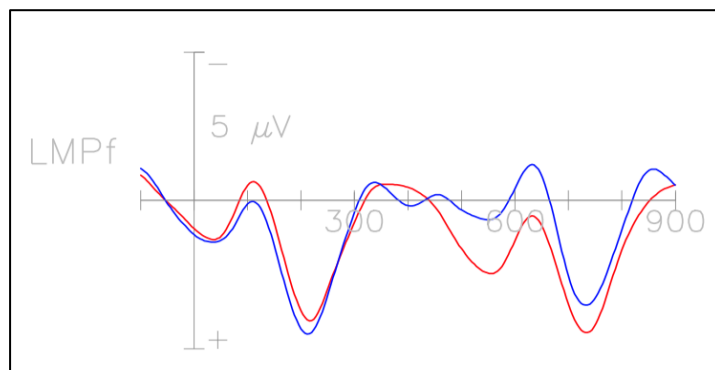
FIGURE 3.7: All participants. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode MiPa. (C) Topographic isovoltage map of the difference between conditions between 500 and 800 msec. Red shading indicates enhanced positivity in response to WCVs.

The results of the distributional ANOVA showed that LAN, N400, and P600 effects were robustly significant in all groups resulting from our ANT-based median split, albeit with quantitative differences in terms of effect magnitude and distribution over the scalp (see TABLE 3.3, supplementary materials for a summary of the results of all comparisons for all groups). In contrast, in the EN time window (see FIGURES 3.8 – 3.13) there were significant condition x anteriority interactions for only two participant groups: the low efficiency orienting ($F[1,39] = 3.78, p = .01$, FIGURE 3.10) and executive groups ($F[1,39] = 2.94, p = .03$, FIGURE 3.12). In both groups, there were larger N100 responses to WCVs than to controls at fronto-central scalp sites. In the low efficiency alerting (Figure 3.8), and high efficiency alerting (Figure 3.9), high efficiency orienting (3.11), and high executive efficiency groups (Figure 3.13), all comparisons in the N100 time window were non-significant (all $p > .29$). These results confirm our hypothesis about the attentional underpinnings of early negativities elicited by WCVs: while all other components differed only in terms of effect magnitude and distribution at the scalp, N100 effects were either present or absent, depending on the orienting and executive efficiency of the participant groups.

(A)



(B)



(C)

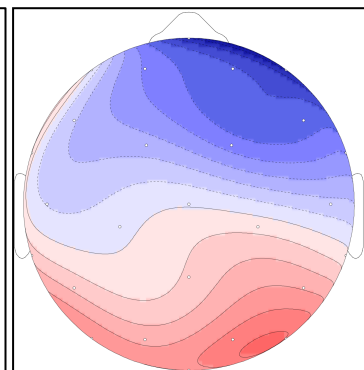
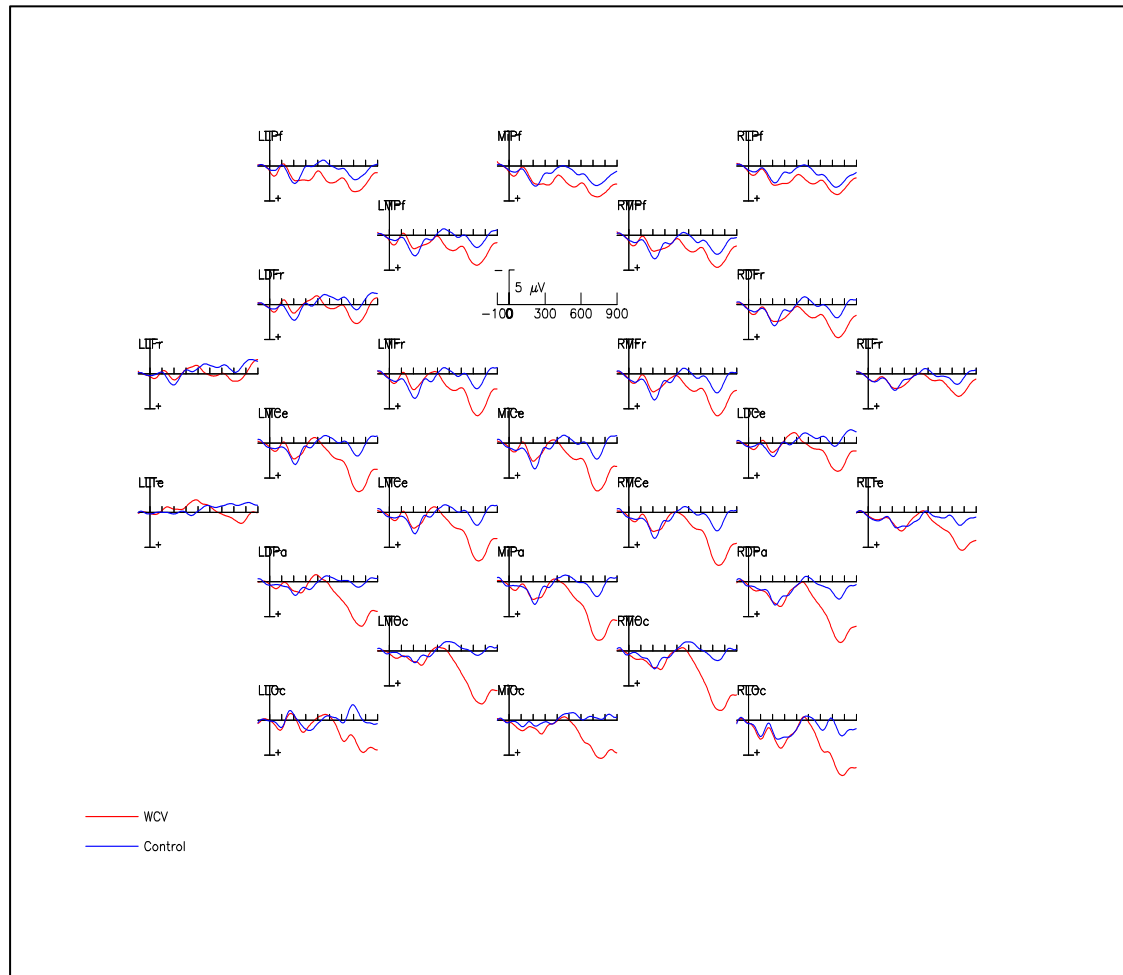
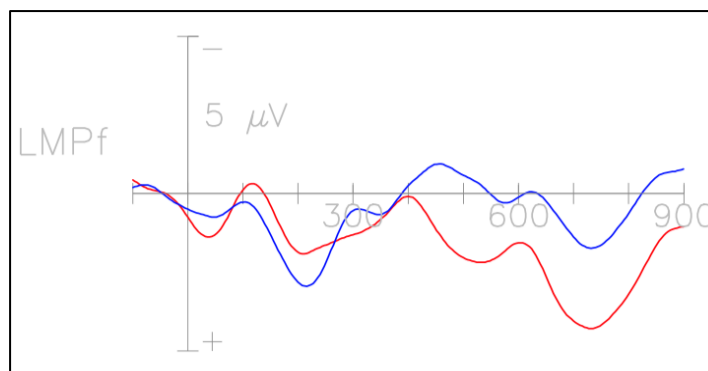


FIGURE 3.8: **Low efficiency alerting group.** (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf, (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec.

(A)



(B)



(C)

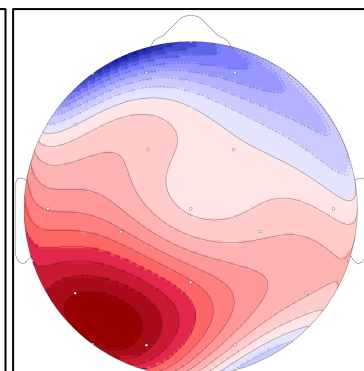
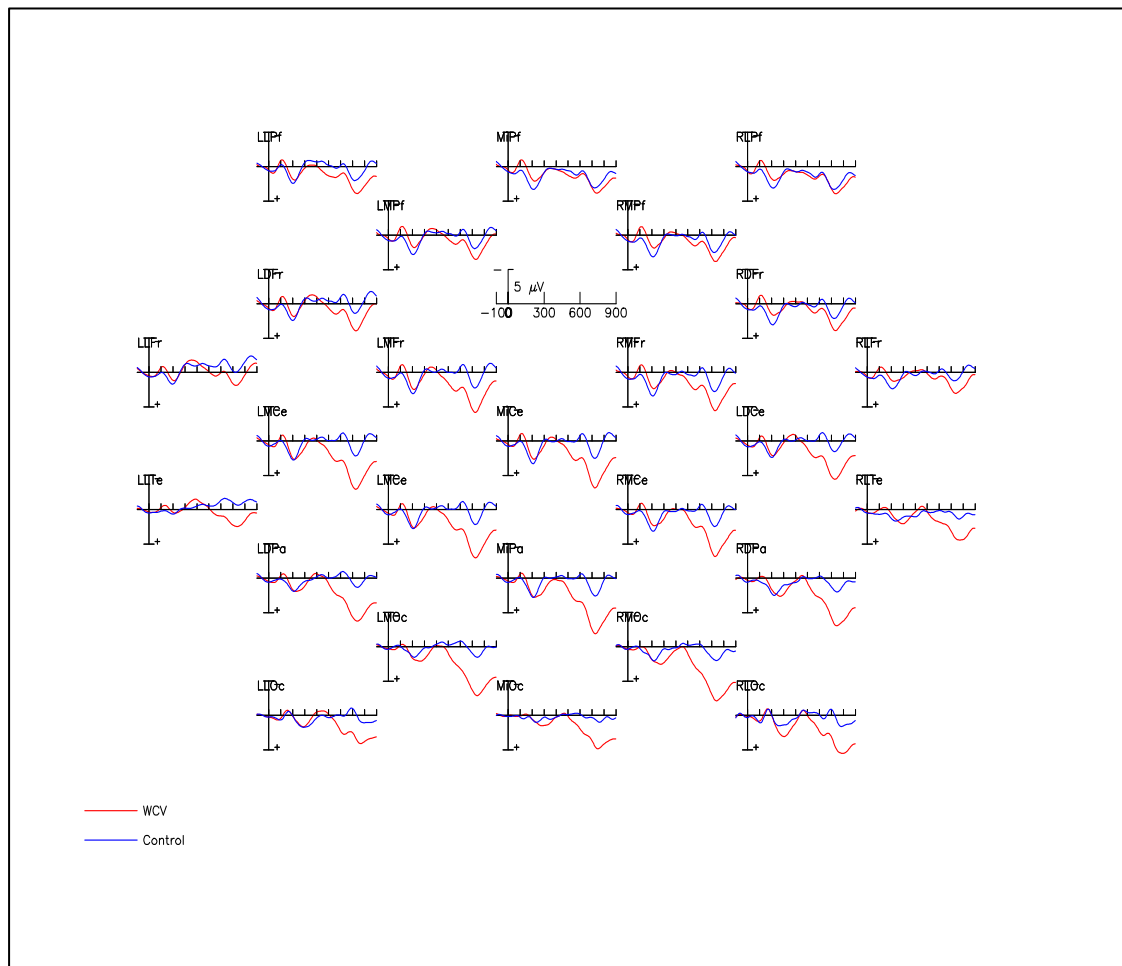
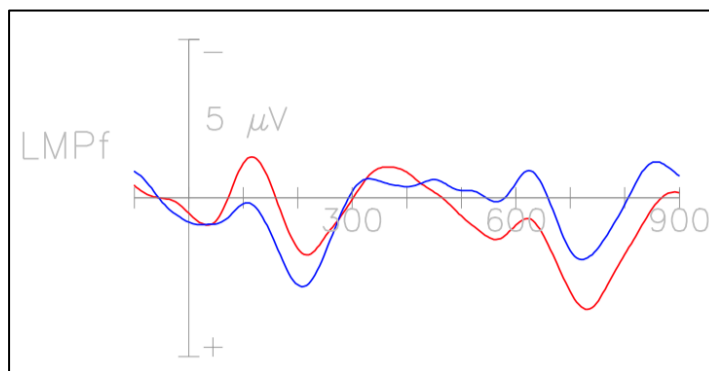


FIGURE 3.9: High efficiency alerting group. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf, (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec.

(A)



(B)



(C)

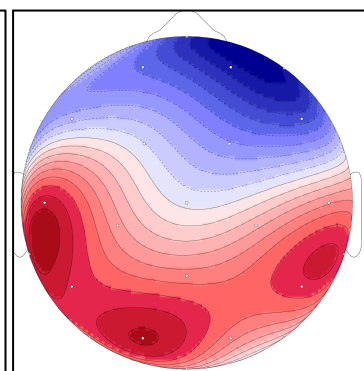
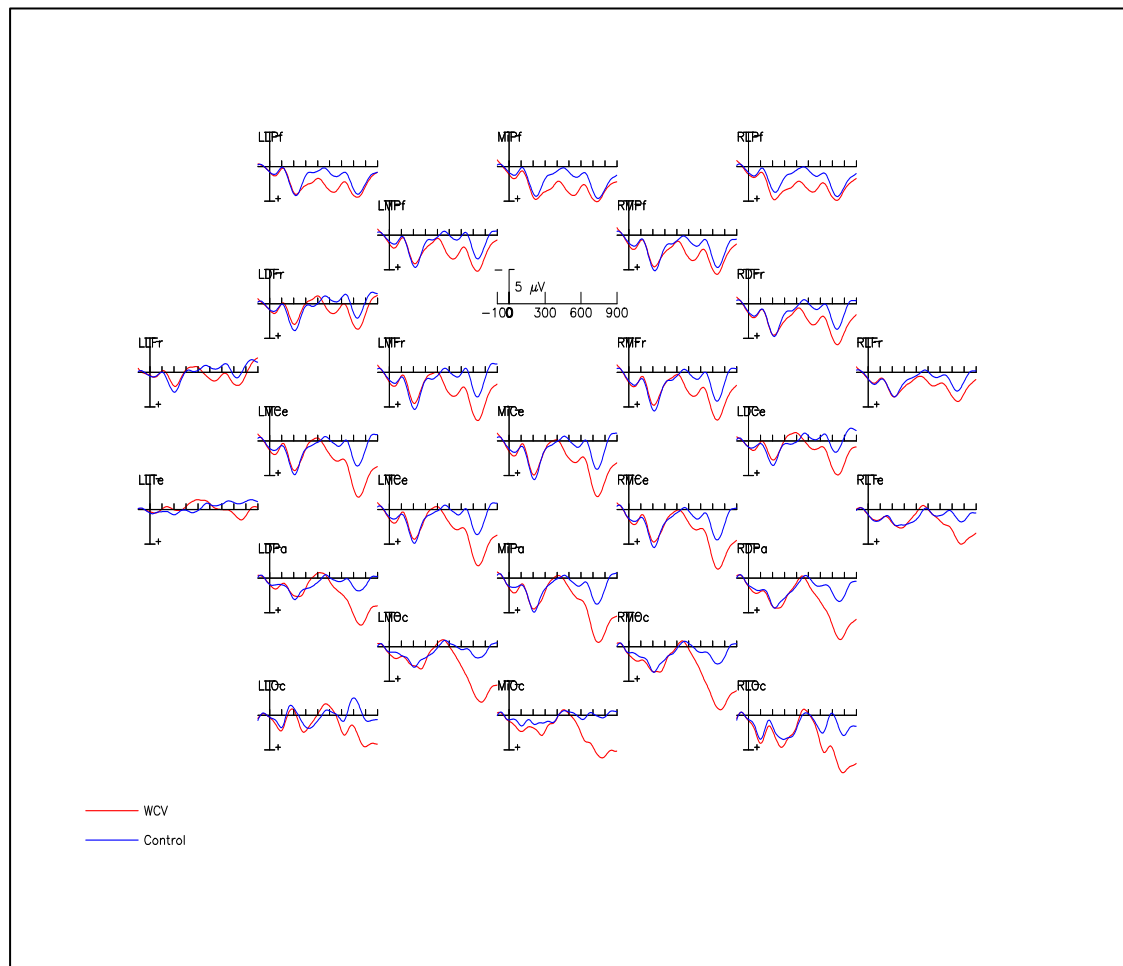
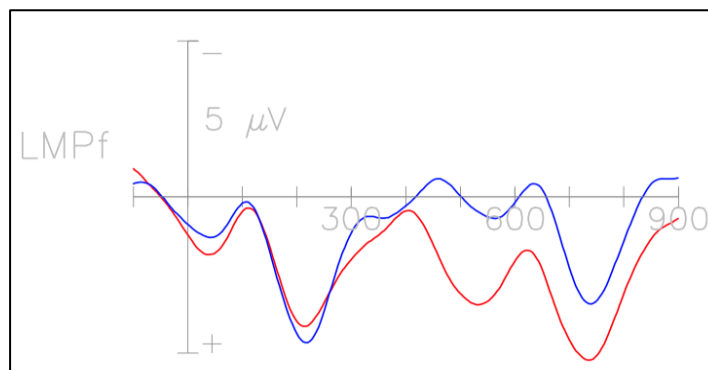


FIGURE 3.10: Low efficiency orienting group. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf. (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec.

(A)



(B)



(C)

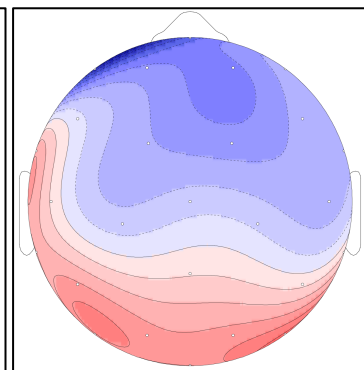
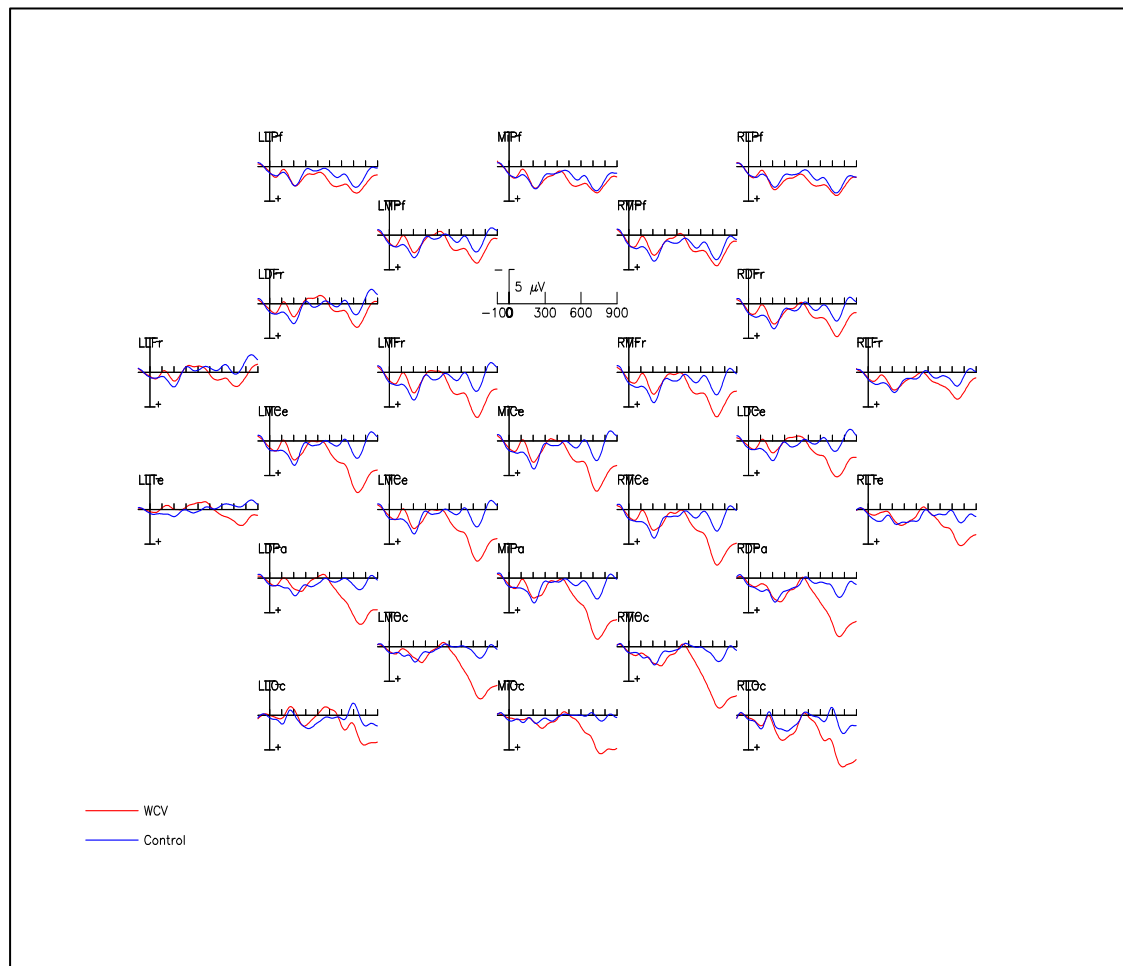
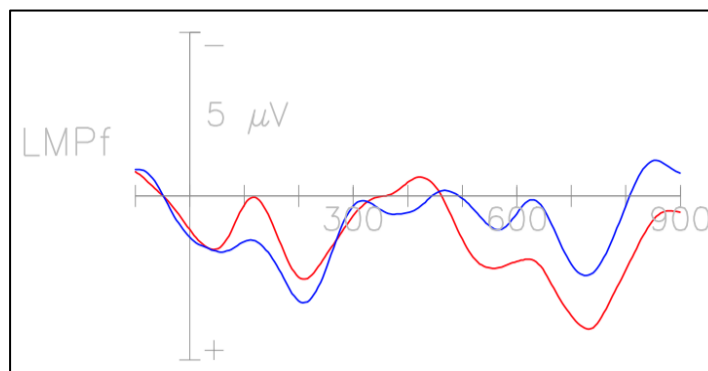


FIGURE 3.11: High efficiency orienting group. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf. (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec..

(A)



(B)



(C)

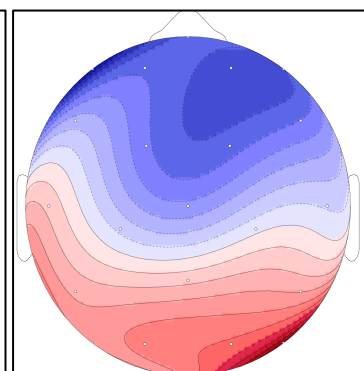
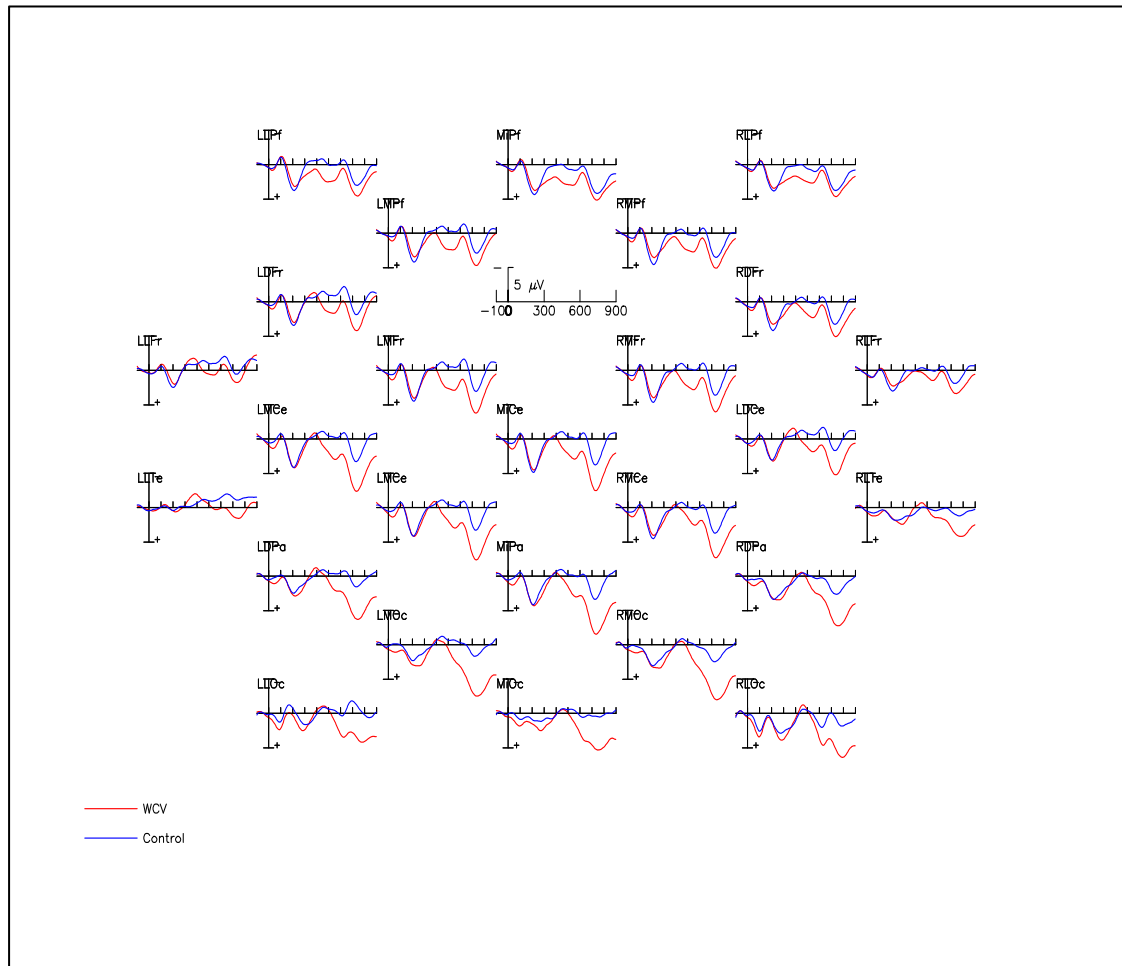
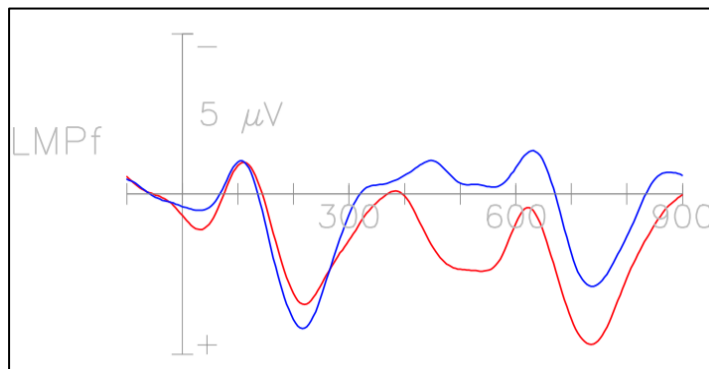


FIGURE 3.12: Low efficiency executive group. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf. (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec.

(A)



(B)



(C)

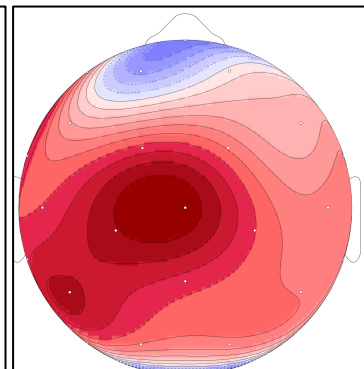


FIGURE 3.13: High efficiency executive group. (A) Grand average ERP waveforms for **WCVs** and **Controls** (1 sec. epoch, 100 msec. baseline). (B) Grand average ERP waveform at electrode LMPf. (C) Topographic isovoltage map of the difference between conditions between 50 and 150 msec.

3.3. Correlational and regression-based analyses

We first looked for simple correlations between our attentional measures and mean N100 amplitude and found no significant relationships (all $R^2 < .2$), indicating that no *single* measure of attentional efficiency was strongly correlated with mean N100 amplitude. We then conducted (multiple) regression-based analyses, comparing models that included all three attentional effect sizes as predictors, as well as the three possible combinations of two attentional effect sizes. Results of these analyses showed that only the regression model containing orienting and executive effect sizes as independent variables significantly predicted mean N100 amplitude. In this model, the combination of orienting ($b = -.11$, $t(36) = 1.91$, $p = .047$) and executive ($b = -.12$, $t(36) = 2.56$, $p = .014$) effect size significantly predicted N100 amplitude at fronto-central scalp sites ($R^2 = .21$, $F(1, 36) = 1.94$, $p = .048$, and $R^2 = .21$, $F(1, 36) = 2.57$, $p = .014$) in a negative linear trend. In other words, the smaller the size of orienting and executive effects, the larger the N100 amplitude, a pattern of results consistent with our median split analysis. Though the amount of variance around mean N100 amplitude accounted for by this model is admittedly small, it is nonetheless significant, as are the predictive contributions of an individual's orienting and executive effect sizes.

4. Discussion

The results of this study provided evidence in support of our hypothesis about the functional significance of early negative responses in sentence processing contexts. After performing median splits on attentional efficiency as assessed by the attention network task (ANT), we found significant LAN, N400, and P600 effects in all of our

participant groups, effects that only differed minimally in effect magnitude and scalp distribution. In contrast, N100 effects were only present in the low efficiency alerting and low efficiency executive function groups. The results of our regression-based analyses provided additional support for our hypothesis, suggesting that while no single measure of attentional efficiency is sufficient to predict N100 amplitude, the combination of an individual's orienting and executive efficiency is able to predict a statistically significant amount of variance in mean N100 amplitude. Below we discuss these findings in the context of the sensory hypothesis (see section 1.2.2), why we believe N100 effects are present only in low-efficiency attention groups, and the domain-general nature of putatively linguistic ERP components.

4.1 Attentional efficiency and the sensory hypothesis

As stated in the introduction (section 1.2.2), the original functional interpretation of early negative responses to word category violations in sentence processing contexts as an index of phrase structure building processes occurring during the first stage of a modular and serial sentence processor (Frazier & Fodor, 1978) has been superseded by a novel view of its functional significance proposed by Dikker and colleagues. In their sensory hypothesis, they argue that early negativities elicited by WCVs actually index physical feature-based sensory operations matching operations. When there is a mismatch between the physical features predicted and those actually encountered, early negativity is elicited. This account has the dual advantage of achieving broad empirical coverage while also explaining the early latency of these effects. However, we felt that the sensory hypothesis lacked an explanation of the feature-predicting mechanism and the reason that feature-based mismatches elicit early negativity. The goal of the current

work was to address these issues, thus adding to the explanatory power of the sensory hypothesis by further elaborating the cognitive processes underlying the elicitation of early negativity in sentence processing contexts.

Our data show that the elicitation of these effects depends on the efficiency of the orienting and executive attention networks. As discussed in section 1.3, the executive attention system functions to control attention in a top-down fashion, constantly generating predictions that modulate activity in sensory cortex. Therefore, this system is a plausible candidate for the mechanism via which the form-based predictions at the core of the sensory hypothesis are generated externally and fed to sensory cortex, continuously modulating expectations about the physical properties of language input. In concert with the actions of the executive attention system, the orienting system functions to shift attention to unexpected but task-relevant input. In the context of sentence processing experiments in which early negativity is elicited, the executive system generates expectations about incoming input, and when the parser encounters a word category violation, the orienting system generates an orienting response to this unexpected but task-relevant input that is inconsistent with previously generated predictions. Under this view, these early negativities are not purely sensory responses, but are more profitably viewed as an attentional modulation of a sensory response, the N100. But why would this orienting response generate an enhanced N100 effect?

As discussed in section 1.4, since Hillyard (1973) it has been well documented that paying selective attention to a stimulus enhances the amplitude of the N100. Additionally, it has been shown that attentional orienting enhances the amplitude of

ERP responses (Mangun & Hillyard, 1987). So, in combination with the attentional mechanisms above, it appears that early negativities elicited in sentence processing contexts are simply selective attention based enhancements of the N100 resulting from interactions between the brain's orienting and executive networks. In sum, we believe that our data provide the sensory hypothesis with additional explanatory power and further clarity on the functional significance of so-called *eLAN* responses. It should be noted however, that our attentional-efficiency based account is based purely on assessing attentional efficiency with the ANT, a task that has previously yet to be employed in concert with a sentence-processing task. Likewise, we only examined the relationship between *visual* sentence processing and a *visual* version of the ANT, and the extent to which the same relationship would hold in the auditory modality remains to be seen.

4.1.1 Why is the N100 effect present only in low efficiency groups?

A valid concern that one might raise is why the N100 effect is elicited only in two out of our six median split groups: low efficiency orienting and low efficiency executive control. Our regression model, which includes both orienting and executive effects, can significantly predict a proportion of mean N100 amplitude, with lower attentional efficiency producing larger effects. While this pattern of results may initially seem counterintuitive, we believe there is a reasonable and simple explanation: it is necessary for those with less efficient neurocognitive systems, in this case the attentional system, to devote additional resources when processing complex language input. For example, studies of the processing of complex language input have shown larger amplitude responses in those participants with low working memory capacity

(see Gunter et al., 2003 and Fiebach et al., 2004 – though cf. also King & Kutas, 1995; Münte et al., 1998; Vos et al., 2001 and Nieuwland & Van Berkum, 2006 for the opposite correlation). Additionally, an fMRI study by Rypma et al. (2002) in which participants performed a working memory task showed that low-span subjects had greater activation in prefrontal cortex than their high-span counterparts, a pattern similar to that observed in an fMRI study by Rypma & D’Esposito (2000). Though the correlational patterns differ across studies, and our account will require further empirical validation, the patterns in our data are consistent with the claim that low-efficiency individuals must devote more (neural) resources to the processing of word category violations, leading to larger N100 responses.

4.2 Domain-general interpretations of ERP effects elicited during language processing

Our data also add to a series of observations suggesting that the notion of ERP responses indexing purely linguistic operations is misguided, and it is rather the case that these responses reflect domain-general cognitive processes operating over linguistic input. Here we have argued that the N100 indexes underlying attentional modulations of sensory responses to the physical properties of linguistic input. Similarly, LAN effects observed during the processing of long-distance distance dependencies and non-canonical word order have been argued to reflect the burdens imposed on working memory by these structures, rather than indexing linguistic operations *per se* (see, among others, Kluender & Kutas, 1993; King & Kutas, 1995; Felser et al., 1997; Münte et al., 2008); Rösler et al, 1998; Matzke et al., 2002; Ueno & Kluender, 2003; Ueno &

Garnsey 2010; Kwon et al., 2013; and Hagiwara et al., 2007). Domain-general accounts have also been proposed for both N400 and P600 effects that have long been interpreted as indexing the semantic and syntactic aspects of language processing, respectively. The N400 is now often interpreted as reflecting the processing of linguistic *and* non-linguistic meaning (see Sitnikova et al., 2008 for a review), with N400 responses elicited, and modulated by, manipulations involving pictures (Ganis et al., 1996), the structure of visual narratives (Cohn et al., 2011; Cohn 2014), videos (Sitnikova et al., 2003), faces (Olivares et al., 1999), environmental sounds (Van Petten & Rheinfelder 1995), gestures (Wu & Coulson, 2005, 2007, 2010), and mathematical sequences (Niedeggen et al., 1999). There has been a long-standing argument about the domain-specificity of the P600 often observed in response to (morpho-)syntactic violations (e.g. Coulson et al., 1998; Osterhout, 1999), a response that is also seen when processing violations of structured musical (Besson & Macar, 1987, Patel et al., 1998) and geometric (Besson & Macar, 1987) sequences.

In conclusion, it appears that brain responses that may initially seem language specific may instead index general cognitive processes such as working memory, meaning construction, and the processing of structured hierarchical sequences, processes that are co-opted by the language system to process linguistic input. Based on the results of the present study, attention can safely be added to this list.

Chapter 3, in full, is currently being prepared for submission for publication. Barkley, C., Kluender, R., & Kutas, M (*in preparation*). The dissertation author was the primary investigator and preparer of this manuscript.

5. Works cited

- Anllo-Vento, L., & Hillyard, S. (1996). Selective Attention to the Color and Direction of Moving Stimuli: Electrophysiological Correlates of Hierarchical Feature Selection. *Perception and Psychophysics*, **58**(2), 191-206.
- Awh, E., & Jonides, J. (1998). Spatial Selective Attention and Spatial Working Memory. In: *The Attentive Brain*, pp 353-380, Cambridge, MA: MIT Press.
- Awh, E., & Jonides, J. (2001). Overlapping Mechanisms of Attention and Spatial Working Memory. *Trends in Cognitive Sciences*, **5**(3), 119-126.
- Baddeley, A. (1993). *Working Memory and Language: Essays in Cognitive Psychology*. England: Lawrence Erlbaum Associates.
- Besson, M., & Macar, F. (1987). An Event-Related Potential Analysis of Incongruity in Music and Other Non-Linguistic Contexts. *Psychophysiology*, **24**, 14-25.
- Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F.E., and Linden, D.E. (2004). Attentional Systems in Target and Distractor Processing: a Combined ERP and fMRI Study. *Neuroimage*, **22**, 530-540.
- Braver, T.S., Barch, D.M., Gray, J.R., Molfeese, D.L., and Snyder, A. (2001). Anterior Cingulate Cortex and Response Conflict: Effects of Frequency, Inhibition and Errors. *Cerebral Cortex*, **11**, 825-836.
- Chomsky, N. *Aspects of the Theory of Syntax*. Cambridge: MIT Press, 1965. Print.
- Cohn, Neil, Ray Jackendoff, Phillip Holcomb, and Gina Kuperberg. (2011). The grammar of visual narratives: Neural evidence for constituent structure in visual narrative comprehension. *Neuropsychologia*, **64**, 63-70.
- Cohn, Neil. (2014). You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science*, **38**(7), 1317-1359.
- Corbetta, M., Kincade, J., Ollinger, J., McAvoy, M., Shulman, G. (2000). Voluntary Orienting is Dissociated from target Detection in Human Posterior Parietal Cortex. *Nature Neuroscience*, **3**(3), 114-127.
- Corbetta, M., & Shulman, G. (2002). Control of Goal-Directed and Stimulus-Driven Attention in the Brain. *Nature Reviews Neuroscience*, **3**(3), 201-215.
- Corbetta, M., Patel, G., & Shulman, G. (2008). The Re-Orienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, **58**(3), 306-324.
- Coulson, S. King, J., & Kutas, M. (1999). ERPs and Domain Specificity: Beating a Straw Horse. *Language and Cognitive Processes*, **13**, 653-672.
- Cowan, N. (1988). Evolving Conceptions of Memory Storage, Selective Attention, and their Mutual Constraints within the Human Information Processing System. *Psychological Bulletin*, **104**, 163-191.
- Cowan, N. (1995). Memory Theories from A to Z. *Contemporary Psychology*, **40**, 552-555. (Review of Theories of Memory, edited by A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris)

- Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge, U.K.: Cambridge University Press.
- Daneman, M. & Carpenter, P (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, **19(4)**, 450-466.
- Dikker, S., Rabagliatti, H., Pylkkänen, L. (2009). Sensitivity to Syntax in Visual Cortex. *Cognition* **110(3)**, 293-321.
- Dikker, S. (2010a). Predicting and Parsing Language in Time and Space. Dissertation NYU, 2010.
- Dikker, S., Rabagliatti, H., Farmer, T., & Pylkkänen. (2010b). Early Occipital Sensitivity to Syntactic category is Based on Form Typicality. *Psychological Science*, **21(5)**, 629-634.
- Dikker, S., & Pylkkänen, L. (2011). Before the N400: Effects of Lexico-Semantic Violations in Visual Cortex. *Brain and Language*, **118(1-2)**, 23-28.
- Downar, J., Crawley, A.P., Mikulis, D.J., and Davis, K.D. (2001). The Effect of Task Relevance on the Cortical Response to Changes in Visual and Auditory Stimuli: an Event-Related fMRI Study. *Neuroimage*, **14**, 1256–1267.
- Downing, C.J. (1988) Expectancy and Visual-Spatial attention: Effects on Vision. *Journal of Experimental Psychology and Human Perception*, **14**, 188-197.
- Eimer, M. (1996). The N2pc Component as an Indicator of Attentional Selectivity. *Electroencephalography and Clinical Neurophysiology*, **99**, 225-234.
- Ericsson, K., & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review*, **102**, 211-245.
- Eriksen, B. A., Eriksen, C. W. (1974). Effects of Noise Letters upon Identification of a Target Letter in a Non-search Task. *Perception and Psychophysics* **16**: 143–149.
- Fan, J., McCandliss, B.D., Sommer, T., Raz, M. & Posner, M.I. (2002) Testing the Efficiency and Independence of Attentional Networks. *Journal of Cognitive Neuroscience*, **14**, 340-347.
- Fan, J., Fossella, J.A., Sommer, T., Wu, Y., & Posner M.I. (2003). Mapping the Genetic Variation of Executive Attention onto Brain Activity. *Proceedings of the National Academy of Sciences USA*, **100(12)**, 7406-7411.
- Fan, J., Mccandliss, B., Fossella, J., Flombaum, J., Posner, M. (2005). The Activation of Attentional Networks. *NeuroImage*, **26(2)**, 471-479.
- Federmeier, K.D., Segal, J.B., Lombrozo, T., and Kutas, M. (2000). Brain Responses to Nouns, Verbs and Class-Ambiguous Words in Context. *Brain* **123**, 2552-2566.
- Fiebach, C., Vos, S.H. & Friederici, A.D. (2004). Neural Correlates of Syntactic Ambiguity in Sentence Comprehension for Low and High Span Readers. *Journal of Cognitive Nueroscience*, **16(9)**, 1562-1575.
- Frazier, L., & Fodor, J.D. (1978). The Sausage Machine: A New Two-stage Parsing Model. *Cognition*, **6(4)**, 291-325.
- Friederici, A.D. (2002). Towards a Neural Basis of Auditory Sentence Processing. *Trends in Cognitive Sciences*, **6(2)**, 78–84.

- Friederici, AD., 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.
- Friederici, AD., von Cramon, Y., & Kotz, S. (1999). Language-related Brain Potentials in Patients with Cortical and Subcortical Left Hemisphere Lesions. *Brain*, **122(6)**, 1033-1047.
- Ganis, G., Kutas, M., & Sereno, M. (1996). The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading, *Journal of Neuroscience*, **8**, 89-106.
- Garavan, H. (1998). Serial Attention within Working Memory. *Memory and Cognition*, **26**, 263-276.
- Gunter, T., Friederici, AD., & Hahne, A. (1999). Brain Responses During Sentence Reading: Visual Input Affects Central Processes. *NeuroReport*, **10(15)**, 3175-3178.
- Gunter, T., Wagner, S., & Friederici, AD. (2003). Working Memory and Lexical Ambiguity Resolution as revealed by ERPs: A Difficult Case for Activation Theories. *Journal of Cognitive Neuroscience*, **15(5)**, 643-657.
- Hahne, A., & Jescheniak. (2001). What’s Left if the Jabberwock Gets the Semantics? An ERP Investigation into Semantic and Syntactic Processes During Auditory Sentence Comprehension. *Cognitive Brain Research*, **11(2)**, 199-212.
- Hahne, A., Eckstein, K., & Friederici, AD. (2004). Brain Signatures of Syntactic and Semantic Processes during Language Development. *Journal of Cognitive Neuroscience*, **16(7)**, 1302-1318.
- Hillyard, S., Hink, R., Schwent, V., & Picton, T. (1973). Electrical Signs of Selective Attention in the Human Brain. *Science*, **182(4108)**, 177-180.
- Hopfinger, J., Luck, S., & Hillyard, S. (2004). Selective Attention: Electrophysiological and Neuromagnetic Studies.
- Indovina, I., and Macaluso, E. (2007). Dissociation of Stimulus Relevance and Saliency Factors During Shifts of Visuospatial Attention. *Cerebral Cortex*, **17**, 1701–1711.
- Kaan, E., Winjen, F., and Swaab, T.Y. (2004). Gapping: Electrophysiological Evidence for Immediate Processing of ‘Missing’ Verbs in Sentence Comprehension. *Brain and Language* **89**, 584-592.
- King, J. & Kutas, M. (1995). Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage During Reading. *Journal of Cognitive Neuroscience*, **7(3)**, 376-395.
- Kirino, E., Belger, A., Goldman-Rakic, P., & McCarthy, G. (2000). Prefrontal Activation Evoked by Infrequent Target and Novel Stimuli in a Visual Target Detection Task: An Event-Related Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, **20(17)**, 6612-6618.
- Lewis, R., Vasishth, S., & Van Dyke, J. (2006). Computational Principles of Working Memory in Sentence Comprehension. *Trends in Cognitive Sciences*, **10(10)**, 447-454.
- Luck, S., Fan, S., & Hillyard, S. (1993). Attention-related Modulation of Sensory-evoked Brain Activity in a Visual Search Task. *Journal of Cognitive Neuroscience*, **5(2)**, 188-195.
- Mangun, G. R., Hillyard, S. A. (1987). The Spatial Allocation of Attention as Indexed by Event-Related Brain Potentials. *Human Factors* **29**, 195-211.

Marois, R., Leung, H.C., and Gore, J.C. (2000). A Stimulus-driven Approach to Object Identity and Location Processing in the Human Brain. *Neuron*, **25**, 717–728.

Marois, R., & Ivanoff, J. (2005). Capacity Limits of Information Processing in the Brain. *Trends in Cognitive Sciences*, **9(9)**, 296-305.

Martinez, A., DiRusso, F., Anllo-Vento, L., Sereno, M., Buxton, R., & Hillyard, S. (2001). Putting Spatial Attention on the Map: Yiming and Localization of Stimulus Selection Processes in Striate and Extrastriate Visual Areas. *Vision Research*, **41(10-11)**, 1437-1457.

Moreno, E., Federmeier, K., and Kutas, M. (2002). Switching Languages, Switching Palabras (Words): An Electrophysiological Study of Code Switching. *Brain and Language* **80**, 188-207.

Müntz, T., Sciltz, S., & Kutas, M. (1998). When Temporal Terms Belie Conceptual Order. *Nature*, **395(6697)**, 71-73.

Näätänen, R., & Michie, P. (1979). Early Selective Attention Effects on the Evoked Potential: A Critical Review and Reinterpretation. *Biological Psychology*, **8(2)**, 81-136.

Näätänen, R. (1982). Processing Negativity: An Evoked-potential Reflection. *Psychological Bulletin*, **92(3)**, 605-640.

Neuhaus, A.H., Opgen-Rhein, C., Urbanek, C., Gross, M., Hahn, E., Ta, T.M.T., Koehler, S., Dettling, M., 2009. Spatiotemporal mapping of sex differences during attentional processing. *Human Brain Mapping*, **30**, 2997–3008.

Neuhaus, A., Urbanek, C., Rhein, C., Hahn, E., Ta, T., Koehler, S., Gross, M., & Dettling, M. (2010). Event-related Potentials Associated with the Attention Network Task. *International Journal of Psychophysiology*, **76(2)**, 72-79.

Neville, H.J., Nicole, J., Barss, A., Forster, K., & Garret, M. (1991). Syntactically-based Sentence Processing Classes: Evidence From Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, **3**, 155-170. **75(4)**, 661-668.

Niedeggen, Rösler, M., Jost, K. (1999). Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. *Psychophysiology*, **36**, 307-324.

Nieuwland, M., & Van Berkum, J. (2006). Individual Differences and Contextual Bias in Pronoun Resolution: Evidence from ERPs. *Brain Research*, **1118(1)**, 155-167.

Oberauer, K. (2002). Access to Information in Working Memory: Exploring the Focus of Attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **28(3)**, 411-421.

Olivares, E., Iglesias, J., & Rodríguez-Holguín, S. (1999). Long-Latency ERPs and Recognition of Facial Identity. *Journal of Cognitive Neuroscience*, **13**, 136-151.

Osterhout, L., & Holcomb, P. (1992). Event-related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language*, **31(6)**, 785-806.

Osterhout, L. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, **14**, 1-14.

- Patel, A., Gibson, T., Ratner, J. & Besson, M. (1998). Processing Syntactic Relations in Language and Music: An Event-Related Potential Study. *Journal of Cognitive Neuroscience*, 10, 717-733.
- Posner, M. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, **32**, 3–25.
- Posner, M., & Peterson, S. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, **13**, 25-42.
- Postle, BR., Awh, E., Jonides, J., Smith, EE., & D’esposito, M. (2004). The Where and How of Attention-based Rehearsal in Spatial Working Memory. *Cognitive Brain Research*, **20**, 194-205.
- Roll, M., Horne, M., and Lindgren, M. (2007). Object Shift and Event-related Brain Potentials. *Journal of Neurolinguistics* **20**, 462-481.
- Rosenfelt, L., Kluender, R., and Kutas, M. (2009). *unpublished manuscript*.
- Ruchkin, D., Grafman, J., Cameron, K., & Berndt, R. (2003). Working Memory Retention Systems: A State of Activated Long-term Memory. *Behavioral and Brain Sciences*, **26(6)**, 709-728.
- Serences, J. T., Schwarzbach, J. & Yantis, S. (2001) Control Mechanisms of Object-based Visual Attention in Human Cortex. *Soc. Neurosci. Abstr.* **27**, 348-349.
- Serences, J., Shomstein, S., Leber, A., Golay, X., Egeth, H., & Yantis, S. (2005). Coordination of Voluntary and Stimulus-driven Attention in Human Cortex. *Psychological Science*, **16(2)**, 114-122.
- Shulman, G. L., McAvoy, M. P., Cowan, M. C., Astafiev, S. V., Tansy, A. P., Corbetta, M. (2003) A Quantitative Analysis of Attention and Detection Signals During Visual Search. *Journal of Neurophysiology*, **90**, 3384-3397.
- Silver, M., Ress, D., & Heeger, D. (2007). Neural Correlates of Sustained Spatial Attention in Human Early Visual Cortex. *Journal of Neurophysiology*, **97(1)**, 229-237.
- Sitnikova, T., Kuperberg, G., & Holcomb, P. (2003). Semantic Integration in Videos of Real-world events: An Electrophysiological Investigation. *Psychophysiology*, 40, 160-164.
- Sitnikova, T., Holcomb, P., & Kiyonaga, K. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, 20, 237-257.
- Smyth, MM., & Scholey, KA. (1994). Interference in Immediate Spatial Memory. *Journal of Memory and Cognition*, **22(1)**, 1-13.
- Streb, J., Henninghausen, E., & Rosler, F (2004). Different Anaphoric Expressions are Investigated by Event-related Potentials. *Journal of Psycholinguistic Research*, **33**, 175-201.
- Petten, C., & Rheinfelder, (1999). Conceptual Relationships between Spoken Words and Environmental Sounds: Event-related Brain Potential Measures. *Neuropsychologia*, **33**, 454-508.
- Vogel, E., & Luck, S. (2000). The Visual N1 Component as an Index of a Discrimination Process. *Psychophysiology*, **37**, 190-203.
- Vogel, E., McCollough, A., & Machiwaza, M. (2005). Neural Measures Reveal Individual Differences in Access to Working Memory. *Nature*, **438**, 500-503.

- Vos, S., Gunter, T., Kolk, H., & Mulder, G. (2001). Working Memory Constraints on Syntactic Processing: An Electrophysiological Investigation. *Psychophysiology*, **38**(1), 41-63.
- Weber-Fox, C., & Neville, H. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, **8**(3), 231-256.
- Wu, Y.C. & Coulson, S. (2005). Meaningful Gestures: Electrophysiological Indices of Iconic Gesture Comprehension. *Psychophysiology* **42**: 654-667.
- Wu, Y.C. & Coulson, S. (2007). How Iconic Gestures Enhance Communication: An ERP Study. *Brain & Language* **101**:234-245.
- Wu, Y., & Coulson, S. (2010). Gestures modulate speech processing early in utterances, *Neureport*, 21, 521-526.
- Zhang, Y. and Zhang, J. (2008). Brain Responses to Agreement Violations of Chinese Grammatical Aspect. *NeuroReport* 19 (10), 1039-1043.

CONCLUSION

Summarizing across the results of the three experiments described herein, I provided support for the central role of domain-general cognitive systems in both language acquisition and the subsequent application of this acquired linguistic knowledge during on-line sentence comprehension. In Chapter 1, I discussed the results of an experiment designed to investigate the role of working memory in the formation of straightforward long-distance antecedent-pronoun relations. The results showed that the second element in such dependencies elicits LAN effects similar to the responses seen at the second element in long-distance syntactic dependencies. These results suggest that, independent of the linguistic level at which a long-distance relationship between two non-adjacent elements is formed, the operations of the working memory system appear to underlie the process. These findings further emphasize the central role that performance factors play in real-time language processing by showing that these factors are crucial to the formation of long-distance relationships at multiple linguistic levels. It appears that the same performance factors appear to be triggered by long-distance linguistic relationships that have been long been argued to constitute distinct aspects of the competence grammar (Reinhart, 1983), (perhaps) reducing the specificity of grammatical knowledge that must be encoded in the competence grammar.

Chapter 2 contained a discussion of the results of an artificial grammar-learning study that investigated the role of selective attention in the implicit learning of linguistic information, showing that selective attention exerts an influence on the behavioral and neural outputs of implicit learning. Additionally, as the brain responses elicited are generally assumed to index general cognitive operations, I argue that, at least during the

earliest stages of language acquisition when proficiency is low, linguistic input is initially acquired using, and processed by, domain-general cognitive systems. Lastly, the results of the experiment described in Chapter 3 show that the elicitation of EN in sentence processing contexts depends at least partially on an individuals' attentional efficiency. This finding adds to an accumulating series of observations suggesting that the labeling of brain responses to language stimuli as "linguistic ERP components" is misguided, and that these brain responses should rather be characterized as indexing general cognitive processes operating on linguistic input. Below, I discuss each experiment in turn.

The results of the experiment described in Chapter 2, in which I compared pronominal and proper name subjects with and without antecedents, showed that pronominal subjects with antecedents elicited LAN effects, similar in onset and distribution to those observed at second elements in long-distance syntactic dependencies. As LAN effects such as these have long been argued to index (some form of) working memory-based operations (see, for example, Kluender & Kutas, 1993; King & Kutas, 1995; Matzke et al. 2002; Gunter et al., 2003; Ueno & Garnsey, 2010; Hagiwara et al., 2007; and Kwon et al., 2013), I argued that the effects observed in the current experiment likewise index working memory-based operations. The burdens imposed by the processing of these long-distance relationships are similar, and therefore the brain responds in kind, meeting these challenges via the operations of the working memory system as indexed by LAN. These findings receive additional support from similar observations in the non-linguistic visual working memory literature [best citation here] While there is a paucity of research pertaining to the issues outlined here

(which has typically focused on encoding, delay-period activity, and fMRI studies attempting to identify the brain areas responsible for these processes rather than the retrieval and integration processes that are the focus of the current work), I feel that by developing non-linguistic paradigms isomorphic to the paradigm used in the experiment described above, it will be possible in the future to further examine the relationships between visual and linguistic working memory, this providing additional support for the arguments forwarded here.

There is a shortage of recent studies that contain non-linguistic manipulations similar to those described above (Fedorenko et al., 2006; Fedorenko et al. 2007; Fedorenko et al., 2009). However, these studies have shown that in a variety of paradigms, linguistic and non-linguistic integration processes in working memory tasks appear to engage similar, modality-independent, working memory processes. As such, these studies provide evidence for domain-general accounts of language processing.

The hypothesis that modality-independent working memory resources operate over any type of structured input is an empirical question, and one that I would like to pursue in future research in order to provide evidence in addition to that outlined above. For example, one could construct experimental paradigms in which participants are sequentially presented with a structured stimulus set (be they faces or colored shapes, for example). If participants were to be trained in advance of this task, for example by telling them that spatially and temporally distal red squares and blue circles constitute a long distance dependency, analogous to those seen in language, and LAN effects were to be observed at the second element (in this case, the blue circle), this would provide more concrete evidence that LAN effects index modality-independent working memory

operations and provide further support for the notion that, regardless of the type of input, the same general working memory processes are triggered by the formation of long-distance relationships between stimuli of any type.

In the experiment described in Chapter 2, I investigated the role of selective attention in the acquisition of grammatical patterns in a traditional training-testing implicit learning paradigm. Participants performed a 2AFC task during brief exposure to an artificial grammar while being explicitly instructed to learn the language's complex agreement system. Unbeknownst to them, the artificial grammar also contained two other sets of (potentially learnable) regularities. Information related to verbal subcategorization, i.e. verb transitivity and the obligatory and lexically-specified constraints on auxiliary verb selection, were encoded at the same matrix verb position that was critical in determining licit agreement relations. Explicit instructions ensured that only the complexities of the agreement system were attended during the training session. During a subsequent testing session, while their EEG was recorded, participants were instructed to judge the grammaticality of grammatical control sentences in the artificial language, as well as agreement violations, and violations of the transitivity and auxiliary selection systems that were unattended during training. Grammaticality judgment accuracy for the two unattended violations was at below chance levels, while accuracy for the control and agreement violations was statistically above chance, evidence that participants had successfully learned this system during the training session. In contrast to their behavioral performance, participants' brain responses successfully discriminated between grammatical sentences and all violation types,

suggesting that some form of learning had in fact taken place even if the result of this learning was not sufficiently robust to be detectable in behavioral measures.

The brain responses observed in this study have all been argued to index general cognitive operations, all being elicited in a variety of paradigms across input types and stimulus modalities. The data showed modulations of the N100, which I argued index processes of attentional capture as participants reach the critical matrix verb position that dictates the grammaticality of the sentences. Subsequent to effects in the N100 time window, there were enhancements of the amplitude of the P200 for all violation types compared to controls. I argued that these amplitude enhancements plausibly reflected processes of feature-based selection of meaningful (but unexpected) and task-relevant stimuli, i.e. violations. Lastly, there were differences in the 300 – 500 msec. (all three violation types again differed from grammatical sentences) and 500-800 msec. time windows, and while we remain agnostic as to the functional identity of these effects, my preferred interpretation is that these effects are representative of a P3b potentially indexing processes of event categorization (Kok, 2001). Under this view, the fact that it was the transitivity violations that elicited enhanced positivity between 500 – 800 msec. may result from the fact that these violations are most difficult to categorize as ungrammatical due to the fact that participants must take the entire sentence into account in order to do so. We feel that these interpretations are at a minimum consistent with the cognitive processes plausibly engaged during the processing of our artificial language and, once more, it is crucial to note that I do not feel that the functional identity of the observed responses is crucial to our claim about the role of selective attention in implicit learning. I simply argue that selective attention influences

implicit language learning, and in doing so modulates ERP components indexing general cognitive operations. For the case of language acquisition, these findings suggest that there is no need to posit the existence of a language acquisition device guided by the (linguistically sophisticated) innate knowledge specified in Universal Grammar (for example in the form of principles and the parameters (Culicover, 1997)), or knowledge of the ordering of the processes necessary for the formation of morphologically complex words (Kiparsky, 1982, Gordon, 1995). Clearly, attentional processes play a crucial role, and if selective attention acts in concert with the learning mechanisms outlined in the introduction, enabling the acquisition of grammatical patterns, this constitutes strong evidence in support of a central role for domain-general performance factors in language acquisition, and reduces the need for highly specified innate linguistic knowledge.

While the interface between the language and working memory system has been extensively investigated, less attention has been paid to the language-attention interface. In Chapter 3, I demonstrated that the elicitation of EN in sentence-processing contexts is dependent on attentional efficiency, with these effects elicited in both low-efficiency orienting groups and those with low-efficiency executive control systems. This pattern was in contrast to other “linguistic ERP effects” (though see Chapter 1), which were statistically robust in all groups resulting from our efficiency-based median splits, and only differed minimally in terms of effect magnitude and distribution over the scalp. In combination with the findings of Dikker (2009) and Dikker et al. (2010, 2011), these findings appear conclusively to refute the original (language-specific) functional interpretation of these effects (Friederici, 1993). If these responses can be interpreted as

indices of attentional modulations of sensory processes, then in no way can they be reducible to underlying processes of initial phrase-structure building operations (Friederici, 1993). Once more, observing modulations of domain-general components (in this case, the N100), our data suggest that domain-general performance factors, operating over structured linguistic input, strongly influence the elicitation of brain responses once thought to be linguistic in nature.

Across three experiments, I have demonstrated the crucial role of general cognitive operations in language acquisition and linguistic processing. It appears to be the case that working memory, selective attention, and attentional orienting and executive efficiency *all* play a crucial role in language acquisition and the use of linguistic competence during on-line language comprehension. As such, these findings suggest that performance factors may well be at the core of these processes, and therefore that these *performance* factors play a more central role in language comprehension than has previously been assumed. In sum, our results suggest that the degree of linguistic knowledge argued to be encoded in a Universal Grammar may be over-specified, and it is fact domain general cognitive systems that are central to guiding both language acquisition and on-line language comprehension. In concert with the powerful learning mechanisms discussed in Chapter 2, mechanisms that take advantage of the rich statistics of the environment and general properties of human perceptual systems, these domain-general processes operating over linguistic representations appear sufficient to explain a large portion of both language acquisition and subsequent language behaviors. While the exact conditions under which these learning mechanisms are engaged is still to be comprehensively mapped, this much is

certain: our data are consistent with the notion that performance factors, acting in concert with some form of linguistic competence, play a central role in real-time language use. No doubt, aspects of linguistic competence are neurally “real” but the end state in which a speaker possesses full knowledge of his/her native language appears to be the result of the actions of domain-general cognitive systems, powerful learning mechanisms, and a Universal Grammar that may contain some form of linguistic competence, but that this pre-specified knowledge may not be as richly specified as had been previously argued. Again, I do not deny the important role of the online application of linguistic competence (nor do I, based on the current results, refute the possibility that *some* form of linguistic competence may be genetically encoded), but rather argue that performance factors and powerful domain-general learning mechanisms can account for a multitude of language behaviors, and therefore that an ongoing research program that investigates these factors has much to tell us about the architecture of the language system.

While we may be genetically endowed with genes that encode for the determination of eye color, why posit the existence of highly-specified innate linguistic knowledge if it isn’t necessary?

Works cited

Culicover, P. *Principles and Parameters*. Oxford: Oxford University Publishing, 1997, Print.

Dikker, S., Rabagliatti, H., Pykkänen, L. (2009). Sensitivity to Syntax in Visual Cortex. *Cognition* **110** (3), 293-321.

Dikker, S., Rabagliatti, H., Farmer, T., & Pykkänen. (2010). Early Occipital Sensitivity to Syntactic category is Based on Form Typicality. *Psychological Science*, **21**(5), 629-634.

- Dikker, S., & Pyllkanen, L. (2011). Before the N400: Effects of Lexico-Semantic Violations in Visual Cortex. *Brain and Language*, **118**(1-2), 23-28.
- Fedorenko, E., Gibson, E. & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific resources. *Journal of Memory and Language*, **54**(4), 541-543.
- Fedorenko, E., Gibson, E. & Rohde, D. (2007). The nature of working memory in linguistic, arithmetic and spatial integration processes. *Journal of Memory and Language*, **56**, 246-249.
- Fedorenko, E., Patel, A., Casasanto, D., Winawer, J. & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory and Cognition*, **37**(1), 1-9.
- Friederici, AD., 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.
- Gunter, T., Wagner, S., & Friederici, A. (2003). Working Memory and Lexical Ambiguity Resolution as Revealed by ERPs: A Difficult Case for Activation Theories. *Journal of Cognitive Neuroscience*, **15**, 643-657.
- Hagiwara, H., Soshi, T., Isihara, M., & Imanaka, K. (2007). A Topographical Study on the Event-Related Potential Correlates of Scrambled Word Order in Japanese Complex Sentences. *Journal of Cognitive Neuroscience*, **19**(2), 175-193.
- King, J., & Kutas, M. (1995). Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, **7**(3), 376-395.
- Kwon, N., Kluender, R., Kutas, M., & Polinsky, M. (2013). Subject/object Processing Asymmetries in Korean Relative Clauses: Evidence from ERP data. *Language*, **89**(3), 537-585.
- Gordon, P. (1985). Level Ordering in Lexical Development. *Cognition*, **2**, 73-93.
- Kiparsky, P. Word Formation and the Lexicon. *Philosophy Papers*.
- Kluender, R., & Kutas, M. (1993a). Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. *Journal of Cognitive Neuroscience*, **5**(2), 196-214.
- Kok, A (2001). On the Utility of the P300 as a Measure of Processing Capacity. *Psychophysiology*, **38**, 557-577.
- Matzke, M., Mai, H., Nager, W., Russeler, J., & Münte, T. (2002). The Costs of Freedom: An ERP Study of Non-canonical Sentences. *Clinical Neurophysiology*, **113**, 844-852.
- McCullough, A., Machizawa, M., & Vogel, E. (2007). Electrophysiological Measures of Maintaining Representations in Visual Working Memory. *Cortex*, **1**, 77-94.
- Reinhart, T. (1983). Coreference and Bound Anaphora: A Restatement of the Anaphora Questions. *Linguistics and Philosophy*, **6**, 47-88.
- Ueno, M. & Garnsey, S. (2007) An ERP Study of the Processing of Subject and Object Relative Clauses in Japanese. *Language and Cognitive Processes*, **23**(5), 646-688.

Ueno, M., & Kluender, R. (2003). Event-related Brain Indices of Japanese Scrambling. *Brain and Language*, **86**, 243-271.