

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Examination of lexical properties during auditory sentence processing using event-related potentials

Permalink

<https://escholarship.org/uc/item/1bg5007n>

Author

Downey, Ryan Andrew

Publication Date

2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Examination of lexical properties during auditory sentence processing
using event-related potentials

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

in

Language and Communicative Disorders

by

Ryan Andrew Downey

Committee in charge:

University of California, San Diego

Professor Tracy Love, Co-Chair
Professor John Moore
Professor John Polich
Professor Frank Haist

San Diego State University

Professor Lewis Shapiro, Chair
Professor Beverly Wulfeck

2006

Copyright

Ryan Andrew Downey, 2006

All rights reserved.

The dissertation of Ryan Andrew Downey is approved, and it is acceptable in
quality and form for publication on microfilm:

Co-Chair

Chair

University of California, San Diego

San Diego State University

2006

DEDICATION

I dedicate this work to my mentor, advisor, and friend, David Swinney. He was the best mentor any student could have hoped for: he pushed, but just hard enough; he encouraged, but not from on high; he insisted, but did so with kindness. I learned much from him about being a thoughtful researcher and making a career out of doing what one loves. He was a truly remarkable person. I was fortunate and honored to have worked in his shadow, for too, too brief a time.

TABLE OF CONTENTS

Signature Page.....	iii
Dedication.....	iv
Table of Contents.....	v
List of Figures.....	vii
List of Tables.....	ix
Acknowledgements.....	x
Vita.....	xii
Abstract.....	xiv
Chapter 1: Introduction.....	1
Chapter 2.....	10
Abstract.....	10
2.1 Introduction.....	11
2.2 Method.....	28
2.3 Results.....	36
2.4 Discussion.....	42
Tables.....	48
Figures.....	49
Appendix A.....	54
References.....	57
Chapter 3.....	61
Abstract.....	61

3.1 Introduction.....	63
3.2 Method.....	78
3.3 Results.....	85
3.4 Discussion.....	89
Tables.....	98
Figures.....	104
Appendix A.....	110
References.....	116
Chapter 4.....	118
Abstract.....	118
4.1 Introduction.....	119
4.2 Method.....	128
4.3 Results.....	134
4.4 Discussion.....	142
Tables.....	151
Figures.....	155
Appendix A.....	166
References.....	173
Chapter 5: Conclusion.....	175

LIST OF FIGURES

Figure 2.1: The Concreteness N400 effect taken from Holcomb et al. (1999)....	48
Figure 2.2: Mean amplitude for abstract and concrete nouns over across both hemispheres in the 300-500 ms window.....	49
Figure 2.3: Grand average ERP waveforms elicited by antecedent nouns rated as being strongly “concrete” or “abstract” in auditorily presented sentences.....	50
Figure 2.4: Grand average ERP waveforms elicited by pronouns co-referent with antecedent nouns.....	51
Figure 2.5: Mean amplitude for abstract pronouns and concrete pronouns over 5 scalp Regions in the 300-500 ms window.....	52
Figure 3.1: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed across point action and process verbs.....	104
Figure 3.2: Grand average ERPs across point action and process verbs when in an enriching or transparent coercion context.....	105
Figure 3.3: Grand average ERPs across enriched and transparent contexts.....	106
Figure 3.4: Grand average ERPs across “ <i>For...</i> ” and <i>at/until</i> modifier types...	107
Figure 3.5: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed.....	108
Figure 3.6: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed, over midline electrodes.....	109
Figure 4.1: Phrase structure I sentences: ERPs elicited by grammatical and ungrammatical sentences in English vs. Jabberwocky.....	155

Figure 4.2: Interaction between language and grammaticality in phrase structure I sentences in the 100-300 ms window.....	156
Figure 4.3: Interaction between language X anteriority in phrase structure I sentences in the 100-300 ms window.....	157
Figure 4.4: Interaction between laterality and anteriority in phrase structure I sentences in the 500-1000 ms window.....	158
Figure 4.5: Phrase structure II sentences: ERPs elicited by grammatical and ungrammatical sentences in English vs. Jabberwocky.....	159
Figure 4.6: Interaction between language X grammaticality X anteriority in phrase structure II sentences in the 100-300 ms window.....	160
Figure 4.7: Interaction between grammaticality X anteriority in English phrase structure II sentences in the 100-300 ms window.....	161
Figure 4.8: Interaction between grammaticality X anteriority in phrase structure II sentences in the 300-500 ms window.....	162
Figure 4.9: Interaction between grammaticality X anteriority in phrase structure II sentences in the 500-1000 ms window.....	163
Figure 4.10: Interaction between phrase structure X language X grammaticality in the 100-300 ms window.....	164
Figure 4.11: Difference ERPs elicited by phrase structure I (black) and phrase structure II (gray) sentences shown over selected electrodes.....	165

LIST OF TABLES

Table 2.1: Attributes of concrete and abstract nouns.....	48
Table 3.1: The 2X2 experimental design crossing factors of Cardinality and Modifier Type used in Todorova et al. (2000).....	98
Table 3.2: Aspectual Coercion Effects: Omnibus Mean Amplitude ANOVAs..	99
Table 3.3: Aspectual Coercion Effects For Verb Type.....	100
Table 3.4: Aspectual Coercion Effects by Modifier Type.....	101
Table 3.5: Aspectual Coercion Effects by Verb Type.....	102
Table 3.6: Aspectual Coercion Effects by Verb Type Only in “For...” Type Modifiers.....	103
Table 4.1: Mean % correct scores for each of the violation types (Phrase Structure I and Phrase Structure II) across all 24 participants.....	151
Table 4.2: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 100-300 Millisecond Latency Range.....	152
Table 4.3: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 300-500 Millisecond Latency Range.....	153
Table 4.4: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 500-1000 Millisecond Latency Range.....	154

ACKNOWLEDGEMENTS

First and foremost, I want to acknowledge my wife, Hollie Suzanne Crower, who has been my primary motivator and supporter since the day I met her. She has gifted me with constant laughter, warmth, partnership, and our beautiful son, Finnegan. These last few months, Hollie has juggled her own busy schedule of teaching, taking classes, and caring for Finn, sacrificing her own rest in order to let me sleep through the night, all so that I would be able to focus on this dissertation. I cannot thank her enough. And Finn, for his part, has been a very good baby; on the days that I watched him, he even let me work for 20 minute stretches before becoming fussy. May he continue taking after his mother.

I would also like to thank John Moore, Beverly Wulfeck, Frank Haist, and John Polich for serving on my committee and for their support and feedback. I am particularly grateful to JP, in whose lab I received much of my initial ERP training, and from whom I learned that the Sword of Science cuts wide and deep.

I would especially like to thank Dr. Lew Shapiro and Dr. Tracy Love, who have mentored me, encouraged me, and gently persuaded me to become more than I would have been otherwise. Lew saw me enter the JDP at the very beginning, and I am extremely grateful and fortunate that he chose to stick with me to the end. His agile thoughts and penetrating mind bring a playfulness to scientific problem solving, and I am constantly motivated by his enthusiasm in tackling issues, particularly those that otherwise seem intractable. Tracy has been an inspiration since I met her; she is thoughtful and long-sighted, with an unbelievably broad perspective and ability to see all sides to every issue, and is able to get more done

in a single day than anyone I have ever known. Meetings with both Lew and Tracy consist of roughly 60% cutting satire of politics, entertainment, and gossip, and 40% intense discussion of science – like chugging psycholinguistics from concentrate. Both have shaped my thoughts and professional aspirations in lasting ways, and have provided me with the guidance, humor, stimulating discussions, and equanimity about 11th hour deadlines that everyone hopes for in an advisor. Most of all, together they stepped in at a time when they themselves were heavily burdened. I will always admire and be grateful for their strength.

Finally, I give special thanks to Dr. David Swinney, in appreciation for everything that he did on my behalf. He believed in me enough to take me under his wing and offer me the incredible gifts of his knowledge, intellect, and respect. He encouraged me to be part of his new ERP lab, an experience which has influenced my career directions profoundly. Every time I left his office, I always said, “Thanks, Dave”, because I had a sense of coming away clearer, more guided, or knowing more than I had when I went in. It became such a reflex that I wonder if he even noticed it as the years passed; I know I scarcely had to think about saying it. And now, there are no more meetings to “talk science”, or hash out future projects, or catch up on personal doings. I realize now how grateful I am for every meeting I had with him. I wish I could remember the last time I thanked him in person. But I can’t. So, here, I give him my continuing, unspoken, but enduring, appreciation.

Thanks, Dave.

VITA

EDUCATION

1994	Bachelor of Science, Psychology University of Michigan, Ann Arbor
1997	Master of Science, Basic Medical Science Wayne State University
2006	Doctor of Philosophy, Language and Communicative Disorders San Diego State University/University of California, San Diego

HONORS AND AWARDS:

“Superior Teaching” Award for Outstanding TA in Cognitive Science Department of Cognitive Science, University of California, San Diego Instructor: Dr. Rafael Nuñez	2006
Pre-Doctoral Fellowship, NIH training grant, “Language, Communication and Brain” Center for Research in Language, University of California, San Diego Mentor: Dr. David Swinney	2004 - 2005
NIDCD Student Fellowship, Research Symposium in Clinical Aphasiology Mentor: Dr. Cynthia Thompson, Northwestern University	2004
ASHA Research in Higher Education Mentorship Program Mentor: Dr. Gloria Waters, Boston University	2000

PUBLICATIONS/POSTERS/ABSTRACTS:

- Wright, HH, **Downey, RA**, Gravier, M, Love, T, & Shapiro, LP (*accepted*). Processing distinct linguistic information types in working memory in aphasia. Aphasiology.
- Gravier, ML, **Downey, RA**, Wright, H, Love, T, & Shapiro, LP (2006). "Processing distinct linguistic information types in working memory in aphasia." Poster presented at the Clinical Aphasiology Conference. Ghent, Belgium.
- Wright, HH, **Downey, RA**, Love, T, Newhoff, M, Schwartz, R, Shapiro, LP (2005). Measuring working memory in aphasia. Journal of International Neuropsychological Society, 11 (S2), 21.
- Wright, HH, **Downey, RA**, Love, T, Newhoff, M, Schwartz, R, Shapiro, LP (2005). "Measuring working memory in aphasia." Poster presented at the International Neuropsychological Society, Dublin, Ireland.
- Downey, RA**, Wright, H, Schwartz, RG, Newhoff, M, Love, T & Shapiro, LP (2004). "Toward a measure of working memory in aphasia." Poster presented at the Clinical Aphasiology Conference. Park City, UT.
- Austermann, SN, **Downey, RA**, Robin, DA, Chapman, SB, & Max, JE (2003). "Sustained attention and discourse performance after childhood stroke." Poster presented at the Clinical Aphasiology Conference. Orcas Island, WA.
- Wright, H, Newhoff, M, **Downey, R** & Austermann, S (2003). Additional data on working memory in aphasia. Journal of the International Neuropsychological Society, 9, 302.
- Wright, HH, Newhoff, M, **Downey, R**, & Austermann, S (2003). "Additional data on working memory in aphasia." Poster presented at the International Neuropsychological Society, Honolulu, HI.
- Downey, RA** & Robin, DA (2001). "Sustained attention in children/adolescents with prenatal and postnatal stroke." Poster presented at the American Speech-Language Hearing Association Conference. New Orleans, LA.
- Boatman, D, Gordon, B, Webber, R, **Downey, R**, Chen, J, Saha, P, Yi, R & Krauss, G (1999). "Contribution of Central Auditory Processing Disorders to variability in Mismatch Negativity." Platform presentation in the Clinical Neurophysiology Symposium at the American Speech Language Hearing Association Conference. San Francisco, CA.
- Müller, R-A, Behen, ME, Muzik, O, Rothermel, RD, **Downey, RA**, Mangener TJ & Chugani, HT (1998). Task-related activations in heterotopic brain malformations: a PET study. NeuroReport, 9, 2527-2533.
- Müller, R-A, Muzik, O, Rothermel, RD, Behen, ME, **Downey, R**, Chakraborty, PK & Chugani, HT (1997). Finger movement in children: less distinct neocortical activation profiles and reduced cerebellar involvement compared to adults. Neuroimage, 5, S11.

ABSTRACT OF THE DISSERTATION

Examination of lexical properties during auditory sentence processing
using event-related potentials

by

Ryan Andrew Downey

Doctor of Philosophy in Language and Communicative Disorders

University of California, San Diego, 2006

San Diego State University, 2006

Professor Lewis P. Shapiro, Chair

Professor Tracy Love, Co-Chair

This dissertation presents three studies using the event-related potential methodology (ERPs) that offer augmentative evidence for investigations of lexical properties during on-line auditory sentence comprehension. Because ERPs are sensitive to specific linguistic properties and have relatively fine-grained temporal resolution, the methodology can be brought to bear on significant issues in psycholinguistic inquiry that have yielded conflicting or inconclusive results. In

particular, ERPs can differentiate between alternatives indistinguishable using behavioral methods.

Experiment 1 examines whether the property of antecedent concreteness is available at the pronoun during co-reference resolution. Results suggest that antecedent concreteness is available relatively immediately at the pronoun in auditory sentences, supporting hypotheses that antecedents are reactivated at coreferent pronouns at a deep, conceptual level.

Experiment 2 investigates the operation of aspectual coercion, in the case where an event that is typically interpreted as a point action event (e.g., “The light flashed at dawn”) is “coerced” into a durative, iterative interpretation by an aspectually marked modifier (e.g., “The light flashed until dawn”). Consistent with linguistically motivated hypotheses, results indicate that aspectual coercion indeed may be a (primarily) semantic operation, manifesting as a lasting negative shift in the ERP in coercion conditions starting immediately when, or shortly after, coercion is licensed. Furthermore, a different effect is found when aspect is “coerced” in process verbs with a continuous but bounded sense (e.g., “The light glowed until dawn”).

Experiment 3 examines the effects of phrase structure violations on the ERP. In this study, participants listened to grammatical and ungrammatical English and Jabberwocky sentences (comprised of semantically null nouns but English verbs and closed class words) to determine whether lexico-semantic influences of verbs properties (e.g., verb subcategorization rules) would be preserved even in language that contains meaningless (semantically null)

arguments. Results indicate that phrase structure violations elicit a P600 effect in both English and Jabberwocky sentences, suggesting increased processing costs in the violation conditions. No LAN to violations was evident in either language for either violation type, contrary to what has been found with identical stimuli presented visually. Results are discussed in the context of prevalent sentence processing models.

Chapter 1

Introduction

The seemingly effortless manner in which most people understand others' speech and language belies the complexity of the task of processing language undertaken by the human brain. The putative mechanisms underlying language processing have been a focus of research for a number of decades. Taking seriously the notion that the brain is an information-processing device, psycholinguistic research has attempted to explain the different information types that are used during language processing, the way such information is activated and integrated, and the time-course of language processing as it unfolds.

This dissertation is concerned with how lexical information – information ostensibly arrayed against a lexical item's entry in the Lexicon, or at least the properties by which a lexical item is associated – is used during sentence comprehension tasks. It is assumed, based on considerable evidence from the literature, that the language processing system can be described by detailing the operations of sub-components such as lexical access, syntactic parsing, discourse processing, and the like. The very fact that the language processing system – in the present case, sentence processing – can be fractionated experimentally requires techniques that can probe the system locally and dynamically as component operations are unfolding over time. Thus, in the studies to be described, neurophysiologic measurements via event-related potentials (ERPs) are used to help characterize the on-going nature of sentence processing.

The dissertation is organized as follows: In this introductory chapter I briefly describe the issues that I will be addressing in the empirical work to follow, and attempt to ground this work in the more general psycholinguistic literature. I take particular care in describing the methodology that is shared among the empirical work. Chapters 2-4 describe the empirical work; Chapter 2 describes an experiment that examines the semantic lexical property of concreteness, and whether this semantic property of an antecedent is reactivated along with the antecedent at a coreferring pronoun. Chapter 3 describes a study examining the process of aspectual coercion, in which a verb's temporal aspect (i.e., whether it describes a one-time or ongoing event) is forced, by a semantic modifier, into a repetitive or continuous interpretation. Chapter 4 describes an experiment designed to test whether syntactic lexical properties – specifically, verb properties related to building phrase structure – are processed normally even when arguments of those verbs are semantically null (e.g., “Jabberwocky” sentences containing pseudowords). Each of these empirical chapters is intended as a “stand alone” example of the psycholinguistic enterprise. The dissertation ends with a General Discussion chapter that takes each experiment and attempts to construct a coherent story about the time-course of lexical and structural processing operations. The discussion also places the results into the context of prominent sentence processing models.

At the extremes of the sentence processing literature lie two accounts. At one extreme are “restricted” accounts, which propose that the process of building a sentence-level representation takes place in stages; in the earliest stages, only structural, or syntactic, information is involved. Such “form-driven” or “syntax-first”

models propose that lexical properties of a combinatorial nature – basic properties such as what part of speech a word is and whether it appears in an appropriate word order for a well-formed sentence, or whether it is marked as a singular or plural – are immediately accessed upon encountering the lexical item; if the properties align with previous context, the word is assigned to the ongoing structure. Lexical properties of a more semantic nature – such as the meaning of different senses of a word or whether the word represents a concrete or abstract concept – are processed later according to this account; conceptual/semantic information is computed only once the item has been properly placed into the accruing structure. Of particular importance to such models is the primacy of structure building before meaning elaboration (e.g., Frazier, 1987; Clifton & Frazier, 1989).

At the other extreme lie “unrestricted” accounts, which propose that each incoming lexical item is immediately analyzed and probabilistically assigned to a structure; based on the most probable combinations of words in a given sentence context, each item’s lexical properties are matched to a probable structure constrained by the individual person’s comprehension abilities, general processing abilities, or experience in the language (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994). Such “interactionist” models hypothesize the interaction between lexical properties (sometimes referred to as “cues”; e.g., MacWhinney & Bates, 1989) which eventually results in a single parse. Thus, in such models, all of a word’s lexical properties are “consulted” initially and immediately, and only the most probable interpretation is constructed. Such models predict that the most probable parse – based on computation and resolution of relevant constraints – is made initially, and if that parse

is determined to be incorrect (again, based on continuous and constant recomputation with each incoming lexical item), then a “reparse” is made. A notable feature of constraint satisfaction accounts is that properties of lexical items determine what syntactic structures (among other things) are appropriate; there is no “external” syntax to which the lexical items are assigned, but rather syntax is computed “on the fly” limited and constrained by individual lexical items. Such accounts have the appeal of being more parsimonious than syntax-first accounts – there is no special privilege assigned, *a priori*, to the processing of syntactic compared to semantic properties. These accounts may be somewhat limited, however, by the computational demands they place on the parser: if each incoming lexical item demands that all possible semantic and syntactic properties are accessed and computed at every moment as comprehension proceeds, the computational demands on the language processing system would be extraordinary.

One strength of “stage” accounts, such as the syntax-first account, is that they provide a less computationally costly mechanism. If only syntactic properties are relevant in the initial stages, an interpretation may be built based only on those properties, independent of semantics, leading to a rapid building of initial structure; this structure can subsequently be reanalyzed if the initial parse turns out to be incorrect. Another strength of “stage” accounts is the inherent “testability” of their predictions. That is, such models predict that the time course of processing syntactic properties (e.g., markings for word class (in languages which have them), tense/aspect, number, and grammatical gender) should be different than that of processing semantic properties (e.g., lexical frequency, thematic properties, animacy, etc.). The empirical

studies presented in this dissertation are based on the assumption that different lexical properties are processed – or, more specifically, manifest their effects – along different time courses during the on-line processing of auditory sentences. Although the present studies were not explicitly designed to test these models, reasonable inferences can be made based on specific results. That is, finding that syntactic properties are processed before semantic properties would offer support for form-based accounts; because form-based accounts indicate that semantics are available relatively later, results showing that semantic properties are immediately available might be taken as support for interactionist accounts.

The first of these studies examines the availability of the lexical property of noun concreteness in the resolution of long-distance dependencies – specifically, when matching a pronoun with the noun to which it refers (i.e., its antecedent). Although a number of investigations have suggested that some semantic properties of an antecedent are reactivated at a coreferring pronoun (Nicol & Swinney, 1989; Love & Swinney, 1995), the precise time course and cognitive substrates of this process remain unclear. This study uses the semantic property of concreteness in an attempt to index the type and time course of lexical information availability during antecedent reactivation at a coreferring pronoun.

Study 2 addresses the phenomenon of semantic enrichment, which poses a unique test for syntax-first models of sentence comprehension: “aspectual coercion” occurs when a lexical item or phrase is “coerced” into generating a different aspectual tense by a governing item in the phrase without any change to the actual syntactic structure of the sentence. For example, “The light flashed at dawn” means that the

light flashed only once, while “The light flashed until dawn” means that the light flashed intermittently for a considerable period of time. The syntactic property of aspect (i.e., whether an event took place once or whether it took place continuously) is altered, not by morphosyntactic features, as is the case in languages with a more rich inflectional system, but by a temporal modifier. Because a different conceptual interpretation is generated based purely on the interaction of lexical items in identical syntactic neighborhoods, the “enriched” interpretation of repetition of the verb – in this case, flashing repetitively – appears to be accomplished by purely semantic means. Indeed, behavioral evidence suggests that the operation of coercion exacts a processing cost, but it is not clear whether this process is of a semantic or syntactic nature. It is also not clear whether the operation of coercion actually begins at the point in time in which it is licensed by the temporal modifier. Study 2 attempts to map out the time-course of this process and determine whether the process of aspectual coercion is manifested semantically on-line.

The third study examines whether the computation of phrase structure (e.g., word class information during sentence construction) is dependent on semantic content. Previous on-line studies have shown the parser¹ to be sensitive to word class information as well as verb subcategorization information (i.e., the syntactic process of matching verbs with their arguments), indicating that these syntactic lexical properties are available and processed almost immediately during structure building. What has not been demonstrated conclusively, however, is whether a verb’s

¹ The parser is the sentence processing mechanism by which humans analyze and ultimately comprehend “who did what to whom”.

subcategorization frames are independent of the semantic properties of the nouns that fill them. Study 3 employs sentences devoid of most semantic information – with the exception of verbs and function words; so-called “Jabberwocky” sentences – to test this question.

All 3 of the current studies use the methodology of ERPs to address these topics. ERPs have two important properties which make them an ideal methodology for these types of investigations: ERPs have a fine-grained temporal resolution, and they have been shown to be sensitive to precisely the kinds of linguistic properties under investigation. Indeed, event-related potentials have been used for over 25 years to investigate a wide variety of psycholinguistic questions. Though the ERP effects relevant to specific questions will be explained in more detail in the respective studies, a (very) brief summary of the methodology is appropriate. For an excellent recent review, the reader is directed to Kutas, Van Petten, and Kluender (2006).

ERPs constitute summated potentials generated in the cortex by bundles of neurons firing in synchrony in response to some cognitive event. ERPs represent voltage measured at the scalp, and are characterized by their amplitude, latency, and scalp distribution. Because of its fine temporal sensitivity and unique ability to differentiate between the processing of different types of linguistic information, the ERP methodology frequently has been used to tap into early processing stages and can provide information regarding the cognitive substrates of language processing.

One of the most thoroughly studied ERP effects in language research is the N400, a negative deflection in the EEG peaking ~400 ms, typically with a centroparietal maximum (often over the right hemisphere), elicited by a cognitive

event such as processing a word, picture, gesture, or sound. The robust N400 was originally found to index a mismatch of semantic expectancy (Kutas & Hillyard, 1980), though it is now generally accepted to be an index of difficulty of integrating an item into a semantic context (Kutas & Hillyard, 1980, 1983; Kutas & Van Petten, 1994); a difference in amplitude between N400s in response to two conditions is referred to as an “N400 effect”. In addition to being able to index the semantic integration of lexical items into an ongoing sentence context, the N400 has been shown to be sensitive to word vs. non-word status (including whether a string of letters constitutes a pronounceable word), word class (nouns, verbs, prepositions, etc.) word length, lexical frequency, word concreteness, and orthographic neighborhood.

Two components commonly used to index syntactic information in psycholinguistic research are the P600/LPC, a positive deflection peaking around 600 ms post-stimulus, and the Left Anterior Negativity (LAN). The P600 is commonly viewed as reflecting a repair or reanalysis process following a syntactic violation or a violation of expected parsing strategy, while the LAN has been shown to index violations of word category, phrase structure violations, and long-distance dependencies such as Wh-movement and gap filling. The ERP methodology thus represents an ideal methodology for investigating questions associated with lexical properties in sentence processing.

To conclude this Introduction, the following dissertation presents 3 relatively independent, stand-alone studies examining the time course and availability of different lexical properties during on-line auditory sentence processing. The

dissertation concludes with a Discussion of how results from the 3 studies provide insight into the nature of the processes underlying comprehension.

Chapter 2

Abstract

An experiment using event-related potentials (ERPs) examines: (1) whether the concreteness effect found with ERPs (greater negativity associated with concrete nouns) by Kounios & Holcomb (1994) may be observed when concrete vs. abstract nouns occur in auditory sentence contexts; and (2) whether the effect (re-)emerges at the point of a subsequent co-referring element, a pronoun. Data from previous reaction time (RT) studies support “deep”, conceptual reactivation of the licensed antecedent (Nicol & Swinney, 1989; Love & Swinney, 1996) at the pronoun; it was therefore predicted that the conceptual property of concreteness in the antecedent would be reactivated at the pronoun. This study used auditory presentation for sentences such as “Phil knows all about prison. He loves to talk about it” versus “Phil knows all about passion. He loves to talk about it” to investigate whether the antecedent’s concreteness, indexed by a concreteness N400 effect (cN400) would be “inherited” by the pronoun. Interestingly, at the antecedents, no concreteness effect appeared. However, a cN400 occurred at the pronouns, with concrete “it” being more negative than abstract “it”, starting ~300 ms following pronoun onset. Taken together, these results provide evidence that deep, conceptual properties of antecedent nouns (concreteness) may be reactivated immediately at the pronoun during auditory presentation of sentences.

2.1. INTRODUCTION

An enduring topic in psycholinguistic research relates to how reference-seeking elements “find”, or link up with, their referents. The use of pronouns to refer to nouns is an obvious example. A native speaker’s ability to process coreference is generally effortless, accurate, and rapid; with the exception of instances of considerable ambiguity, the “on-line” (i.e., moment by moment, unfolding in time) computation of coreference proceeds without hindrance. However, it is not clear when, and according to what time course, different types of information are utilized or engaged in order to yield successful comprehension.

This paper describes an experiment that examines how much of one specific aspect of an antecedent’s conceptual representation is reactivated at a coreferent pronoun during on-line sentence processing. Processing associated with such conceptual information – in this case a noun’s “concreteness” – can also provide information regarding at what point in time that information is available.² For example, the word “prison” in (1a) is considered concrete, while the word “passion” in (1b) is considered abstract. In these sentences, “it” is a pronoun which corefers with either a concrete (1a) or abstract (1b) antecedent.

(1a) Phil knows all about prison. He loves to talk about it.

(1b) Phil knows all about passion. He loves to talk about it.

This experiment was designed to determine the time course of informational availability of semantic lexical property of concreteness, i.e., whether the conceptual

² These questions must be distinguished from questions of when coreference is actually computed, which is the subject of much debate, as described below.

information necessary for coreference computation is available in a timely, on-line manner. Previous studies have examined coreference using other properties of antecedents (see review below), but concreteness permits a unique examination of coreference processing because it indexes the conceptual, semantic properties of the lexical item. In other words, previous examinations of coreference using structural properties such as antecedent number, syntactic gender, and lexical frequency yield evidence that structural properties of the antecedent and pronoun must be matched first before the semantic content of the antecedent is appreciated. A lexical item's concreteness indexes the conceptual representation of the word it refers to, and so evidence that an antecedent's concreteness is available at a pronoun is strongly suggestive that the entire conceptual representation of the antecedent is also available. This issue is addressed in the specific case of pronominal coreference using the ungendered, inanimate, definite pronoun, "it". Thus, in (1a) and (1b) above, even though both items conclude with "it", one "it" is coreferenced to a concrete noun ("prison") while the other coreferences an abstract noun ("passion"). Previous studies (Nicol & Swinney, 1989; Love & Swinney, 1996) suggest that conceptual properties of the antecedent are available immediately at the pronoun. Given that the concreteness of the antecedent can be indexed using ERPs (e.g., Kounios & Holcomb, 1994), evidence that an antecedent's concreteness can be indexed at its coreferent pronoun would suggest that semantic properties are available very early in the process of establishing long-distance structural dependencies.

After a brief review of relevant concepts, an experiment is presented which tests the hypothesis that, immediately upon processing a pronoun in a sentence, the

concreteness of an antecedent noun is available immediately, suggesting a “deep”, conceptual reactivation of the antecedent and its semantic properties.

2.1.1. Pronominal coreference: General Considerations

A substantial linguistic literature attempts to describe how the resolution of pronominal coreference transpires, focusing on structural or syntactic mechanisms as their emphasis.³ For example, Chomsky’s Government and Binding (GB; 1981) theory posits Principle B (among others), which states, roughly, that a non-reflexive personal pronoun must refer to a Noun Phrase (NP) outside of the local clause. Thus, Principle B uniquely specifies the antecedent in examples such as Sentence (2):

- (2) Frodo_i knew that Sam_j wanted to join him_{i/j*} in the hopeless quest.⁴

Principle B dictates that the object “him” must refer to an NP outside of the local relative clause, leaving “Frodo” as the only appropriate coreferent antecedent.

However, certain contexts can be described which require disambiguation. For example, Principle B has difficulty unambiguously assigning an antecedent for “him” in (3) (taken from Nicol, 1988).

- (3) The boxer_i told the skier_j that the doctor_k for the team would blame **him**_{i/j/k*} for the recent injury.

Principle B states that “him” may not co-refer with “doctor”; it is unclear, structurally speaking, whether the “boxer” or “skier” is the one who will be blamed.

³ For a review of a more “functionalist grammar” view of pronoun coreference, see Harris & Bates (2002).

⁴ In this paper, conventional linguistic notation is used to denote co-referents (subscripts) and un-grammaticality (*)

Recent psycholinguistic models have attempted to explain how pronoun assignment is accomplished in such cases, using the concept of antecedent “accessibility”; that is, listeners “prefer” (i.e., are more reliable or quicker at) assigning coreference to antecedents that are more accessible to the parser, compared to competitor antecedents that are, for syntactic, semantic, or pragmatic reasons, less accessible. For example, Gernsbacher and colleagues consistently find facilitation in probe verification tasks in which readers more rapidly and accurately assign coreference to antecedents that were first-mentioned in the preceding context (Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves, & Beeman, 1989), as well as those that were mentioned more recently (Gernsbacher et al., 1989). Matthews and Chodorow (1988) examined the effects of distance between antecedents and pronouns, as well as depth of embedding of the antecedent in structurally embedded sentences, and concluded that more recently available (and less deeply embedded) antecedents are processed more readily. Linguists’ observations that listeners/readers appear more likely to assign coreference to an antecedent that is in the same syntactic and/or thematic position as the pronoun – so-called Parallelism of structure – has found support in the psycholinguistic literature (Gordon & Searce, 1995; Stevenson, Nelson, & Stenning, 1995; and many others). Still other studies have suggested that readers are faster when assigning coreference to antecedents that are subjects, or the “focus” – or, defined differently, topics – of the preceding discourse than those without that status (Gordon, Grosz, & Gilliom, 1993; Arnold, Eisenband, Brown-Schmidt, and Trueswell, 2000; Cowles, 2003). Recently, Foraker and McElree (2006) found that “prominence” of an antecedent, achieved by syntactic

clefting, results in a more available trace of the antecedent in working memory (but does not lead to improved reaction times to identify those antecedents). Finally, research by Badeker and Straub (2001) find evidence for a complex interaction of several factors in determining which antecedents are likely candidates during pronominal coreference. Their results suggest that pronominal resolution may be accomplished via an “interactive-parallel-constraint” model in which numerous weighted constraints (structural binding, prominence in focus structure, and gender-number compatibility) compete to yield the final parse. All of these models share elements of the availability of candidate antecedents, but they come short of describing the psychological operations that occur during the process of antecedent reactivation.

2.1.2. Pronominal coreference: Lexical Properties

Clearly, the parser’s computation of pronoun coreference must involve availability, or “accessibility” constraints such as those listed above, to facilitate access to candidate antecedents. However, accessibility constraints are not the only influences on antecedent assignment, as hinted in the results from Badeker and Straub’s (2001) study. For the most part, the aforementioned models express the constraints on pronoun coreference as a matching process between syntactic features of the pronoun and features of candidate antecedents. Take, for example, the lexical property of animacy. The inanimate “it” results in the barring of non-animate antecedent competitors, as in (4) below.

- (4) Frodo_i showed Sam_j the ring_k, saying they needed to destroy it_{i*/j*/k}

Because “ring” is an inanimate entity, “Frodo” and “Sam” are dispreferred as antecedents for “it”. Indeed, considerable evidence supports such a “matching” process between the lexical properties of the pronoun with candidate antecedents (Nicol, 1988; Nicol & Swinney, 1989; Carreiras, Garnham, & Oakhill, 1993; Garnham & Oakhill, 1985; Garnham, Oakhill, Ehrlich, & Carreiras, 1995; Gernsbacher, 1989).

The lexical features most studied in investigating constraints on pronominal coreference are semantic gender, syntactic gender (in languages which use it), animacy, and number. Although the time course of its effect is debated, implicit or explicit semantic gender (i.e., stereotypical gender associations for roles, professions, and names; e.g., “Fred” being a male name vs. “Ginger” being female, or ballerina being female) has been repeatedly demonstrated to be immediately considered by the parser (Marslen-Wilson, Tyler, & Koster (1993); Carreiras, Garnham, Oakhill, & Cain, 1996; Reynolds, Garnham, & Oakhill, 2006; Rigalleau & Caplan, 2000; Kennison & Trofe, 2003; MacDonald & MacWhinney, 1995; but cf. Greene, McKoon, & Ratcliff, 1992). An antecedent’s animacy (i.e., alive vs. inert; e.g., capable of performing an action vs. limited to being acted upon) can likewise license it or remove it from consideration (Clark & Begun, 1971). Number marking and syntactic gender are similarly well elaborated in the pronoun literature (Nicol, 1988; Nicol & Swinney, 1989; De Vincenzi, 1999; Carreiras, Garnham, Oakhill, & Cain, 1996; Garnham & Oakhill, 1985; Cacciari, Carreiras, & Cionini, 1997; MacDonald & MacWhinney, 1995; Matthews & Chodorow, 1988; Shillcock, 1982; Arnold, et al., 2000).

2.1.3. Pronominal coreference: Antecedent Reactivation

Lexical properties of antecedents themselves – independent of where they are in the discourse and how they are being used, structurally or thematically – thus appear to have significant impact on coreference resolution. Once a candidate antecedent is isolated, however, the degree to which its conceptual properties are available remains controversial. Some authors claim that the antecedent is reactivated in coreference resolution in a complete way (e.g., from the lexical form on up; called the full access hypothesis by van Gompel and Majid, 2004), while other studies suggest that the antecedent is reactivated only at a “deep”, or conceptual, level.

Research by Nicol and others (Nicol, 1988; Nicol & Swinney, 1989; Nicol & Swinney, 2003; Love & Swinney, 1996), for example, supports the latter. Their research suggests that a (structurally licensed) antecedent is reactivated at a downstream anaphor, but that this representation is not at a surface, or “form”, level. In these studies, reactivation of the antecedent has been demonstrated using Cross Modal Lexical Priming (CMLP). In CMLP, participants listen to sentences presented aurally. At some (experimenter-controlled) point during the sentence, a letter string (“probe”) appears on a computer screen, and participants are required to decide whether the letter string constitutes a real word in English while attempting to continue to process the auditory sentence. Some of the target letter strings are actually semantic associates (“related” probes) of candidate antecedents, while other strings are not semantically related to anything in the sentence (“control” probes)⁵. The reaction

⁵ A key manipulation in this comparison is that the baseline RTs to related and control probes are matched *a priori*.

time (RT) measures to related probes are compared to RT to control probes presented at the same time points during sentence, e.g., immediately following the pronoun. Relatively shorter RT to the related probe compared to the control probe is considered evidence that the antecedent – and its semantic associates – were reactivated at that point in the sentence.

Using this methodology, Nicol & Swinney (1989) demonstrated that in a sentence such as (5) below, semantic associates to both “boxer” and “skier” are responded to faster than unrelated control probes (indicating antecedent reactivation) immediately following the pronoun, “him” (probe presentation is indicated by ^).

(5) The boxer told the skier that the doctor for the team would blame him^
for the recent injury.

Compellingly, Nicol and Swinney found that both possible antecedents were reactivated, while “doctor”, which is an inappropriate antecedent (due to Principle B, above), was not.

Furthermore, upon reaching a pronoun, the parser does not appear to activate the antecedent at the basic word form level, as it appears to do during general lexical activation (Swinney, 1979). Love and Swinney (1995) demonstrated that the reactivation of antecedents at pronouns is accomplished by the parser differently than when antecedents are themselves processed, insofar as only the contextually appropriate meaning of lexically ambiguous antecedents were reactivated at the pronoun. For example:

(6) Jeff had read about the problems with savings and loan institutions, so, he went to his **bank**_i¹ to ask about the safety² that *it*_i³ provides with respect to CD investments.

In sentence (6), the word “bank” has more than one possible meaning (e.g., financial institution or side of a river). Probe words related to both of those meanings were presented at each of the probe points above (shown in superscript). If pronominal coreference involves reactivation of the surface, or word, form representation of “bank”, then both its possible meanings should be reactivated (Swinney, 1979) at probe point 3. Love and Swinney (1995) found priming for both meanings of bank at probe point 1, for neither meaning at probe point 2, and for only the contextually appropriate meaning (e.g., facilitation to money but not to river) at the pronoun “it”.⁶ Thus, antecedent reactivation does not appear to reactivate a phonological, surface word form, but rather reactivates a deeper, conceptual representation of the antecedent.

This conclusion has been disputed, however. Schmitt, Meyer, and Levelt (1999) found that the phonological form of the antecedent was activated during pronoun production in German, suggesting full lexical reactivation of the antecedent (though this involved language production, not comprehension). Reading time evidence (van Gompel & Majid, 2004) suggests that an antecedent’s lexical frequency can be indexed at a downstream pronoun. Along these same lines, Heine et al. (2006)

⁶ Crucially, Love & Swinney found no priming at probe point 2, demonstrating that both representations for “bank” had faded prior to the pronoun. The priming at probe point 3 indicated conclusively that only the appropriate sense of “bank” was, indeed, re-activated at the pronoun.

found that ERPs elicited in response pronouns coreferent with relatively infrequent antecedents were associated with greater amplitudes in the N400 and P300 components than amplitudes to pronouns coreferent with more frequent antecedents. This difference in amplitude was graded across three levels of frequency. Because a word's frequency is typically considered to be a property of the "word-form", or lexical entry (cf. Jescheniak & Levelt, 1994), Heine and colleagues suggested that their results support the (surface) reactivation of the antecedents' lexical entry.

Conflicting results such as these must be viewed with an eye toward the specific methodologies used; any methodology or task has the tendency to impose a certain degree of bias onto the data based on the nature of the task (Swinney, 2000; Love & Swinney, 1996; Nicol & Swinney, 2003; Hahne & Friederici, 1999; Nicol, Swinney, Love & Hald, 1997).⁷ Furthermore, RT studies may not have a fine enough temporal resolution to tease apart the series of steps of cognitive processes that must be undertaken during such complex psycholinguistic tasks; e.g., once the linguistic information has been processed, a decision must be formulated and then executed, involving decision making stages, motor planning and movement stages, and so on. This effectively places an upper bound on the ability to detect when the processing in question took place (e.g., when conceptual properties of an antecedent are actually available). However, methods like event-related potentials offer a possibility to

⁷ It is worth noting that a large portion of the observed between-study variability in many psycholinguistic domains can be explained by critically examining the methodologies used (e.g., Nicol, Fodor & Swinney, 1994; Swinney, 2000; Love, 1996).

investigate these on-going, continuous processes with a fine-grained temporal resolution.

2.1.4. Electrophysiology

Ideally, an experimental technique intended to measure the availability of lexical properties associated with coreference processing would be sensitive enough to characterize the processing as it happens (i.e., at the stage of perceptual processing that occurs after the low-level sensory processing but prior to decision making or response selection). Electrophysiology is a methodology that has such fine-grained temporal resolution. Event-related potentials (ERPs) represent the measurement of cortical potentials generated in response to specific cognitive, experimenter-controlled stimuli. Specific patterns and components of the waveforms generated provide information about the processing being undertaken by the brain, with a temporal resolution down to milliseconds (Rugg & Coles, 1995). A particularly valuable aspect of ERPs for present purposes is that ERPs have been shown to be sensitive to different types of linguistic information, including syntactic and semantic information, resulting in distinct markers in the ERP waveforms (cf., Kutas & Hillyard, 1980; Holcomb & Osterhout, 1992; Münte, Heinze & Mangun, 1993).

For example, the N400 component is a negative deflection in the cortical voltage, having a predominantly centroparietal (and often rightward) distribution. It generally peaks between 300 and 500 ms after presentation of a semantically anomalous, unexpected, or low frequency word or event (e.g., Kutas & Hillyard, 1980, 1984) and has been characterized as a manifestation of difficulty in integrating

semantic information into an existing context (Van Petten & Kutas, 1991). The N400 has been shown to be sensitive to manipulations in cloze probability, word frequency, word class (open- vs. closed-), word category (noun, verb, etc.), word length, and the operation of long-distance dependencies (Kutas & Hillyard, 1980, 1983, 1984; Van Petten & Kutas, 1990, 1991; Osterhout, Allen & McLaughlin, 2002; Federmeier & Kutas, 1999; Kluender & Kutas, 1993; Brown, Hagoort & Ker Teurs, 1999; Münte et al., 2001; Hinojosa, Martin-Loeches & Rubia, 2001; and many others).

2.1.5. Concreteness

A word's "concreteness"⁸ is a lexical property associated with the conceptual representation of a noun, and as such has the potential to be indexed by a coreferent pronoun during the reactivation process. Concreteness ratings are typically obtained by asking language users to rate how perceivable a word's referent is through the senses (Paivio, 1986). A word representing a real-world object like "table" tends to be rated as being rather concrete (6.30 out of 7 on Toglia and Battig, 1978, with a 7 being "very concrete"), while a word representing a psychological or esoteric concept like "justice" is rated as being abstract (2.61 out of 7).⁹ Despite their seemingly subjective

⁸ Although there is some debate as to the distinction between "concreteness" and "imageability" (cf., Tyler, Moss, Galpin & Voice, 2002), they are considered to be highly correlated (e.g., $r = 0.83$ in Paivio et al., 1968). Nevertheless, to avoid controversy, "concreteness" in this paper is understood in the same sense as it is employed by Kounios & Holcomb (1994), who had participants rate words as to whether they represented a "concrete object" or an "abstract concept."

⁹ Of course, the particular numbers involved in these rating scales are somewhat arbitrary; concreteness ratings are made on a 7-point scale in Toglia & Battig, 1978, and the Toronto Word Pool, but Coltheart's MRC Database (1981) uses a rating scale from 100 to 700. Regardless of the actual scale used, the meaningful distinction here

nature (i.e., the fact that concreteness ratings are derived from native speakers' intuitions), there is generally strong agreement between raters as to a word's concreteness, as well as reliable relative concreteness/abstractness ratings for the same items across rating systems.

The concreteness effect – in which words rated as being concrete are responded to, named, and represented more quickly than abstract words – has been demonstrated across a number of behavioral, experimental modalities. Two models were originally proposed to account for the findings: the dual-coding model, and the context availability model. The dual-coding model (Paivio, 1986, 1991) claims that abstract and concrete words are represented differently in the brain, using two separate systems, one in each hemisphere. Abstract concepts, which tend to have little or no visual or imagistic features, are processed primarily in the left hemisphere's verbal system. Concrete concepts, on the other hand, have strong visual associations, and therefore use both the left hemisphere verbal system and a system supported by the right hemisphere: the imagistic system. Concrete words are thus more strongly represented in the brain than abstract words, as concrete words are represented in both systems. This view accounts for improved performance for concrete compared to abstract words on memory and naming tasks (Paller, 1990; Bleasdale, 1987), and faster associate generation (de Groot, 1989). The context availability model, on the other hand, has been used to explain why the concreteness effect appears to go away in supportive contexts (i.e., in a sentence as opposed to single word reading).

is the relative concreteness rating obtained across two words (one more concrete and the other more abstract).

According to this model, abstract words are inherently more vague or ambiguous, and they therefore require more effort and time to process than concrete words, leading to the concreteness effect. If a supportive context is built, however, the advantage held by concrete words disappears (Kieras, 1978; Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Stowe, 1989; for a review, see Schwanenflugel, 1991).¹⁰

A series of studies by Holcomb and colleagues has examined noun concreteness in English using ERPs in an attempt to find support for either the dual-coding or context-dependent models. Kounios and Holcomb (1994) investigated whether concrete and abstract words had different neural substrates, comparing the morphology and lateralization of the ERP signature to both types of words. Kounios and Holcomb examined the ERPs resulting from processing concrete words and abstract words in isolation using a lexical decision task (Experiment 1), and a concreteness judgment task within a sentence context (Experiment 2). Results indicated a difference in the N400 component to abstract and concrete words; this effect manifested as a greater negativity for concrete words than abstract words in the 300-500 ms window, with a more anterior distribution than the typical N400. The authors labeled this the “concreteness N400 effect.” In Experiment 2, Kounios and Holcomb (1994) found an even greater concreteness N400 effect between concrete and abstract words in sentence contexts.

Holcomb, Kounios, Anderson, and West (1999) examined the concreteness N400 effect in sentence-final contexts, manipulating degree of violation (congruous

¹⁰ Holcomb and colleagues have recently posed a third model, which is a fusion between the two models, called the “extended context availability” model. This is treated in more detail in the Discussion, below.

versus neutral *versus* incongruous sentence-final words), as well as concreteness.

They found that, again, concreteness showed a significant effect on the N400, with N400s being more negative to concrete words than to abstract words. This concreteness N400 effect was especially marked in the incongruous sentences, but it disappeared in the congruous and neutral contexts, suggesting that supporting context may reduce the effect. Thus, in both studies, the authors found that nouns classified as abstract and concrete demonstrated differential ERP patterns, with ERPs to concrete nouns being more negative than abstract, particularly over anterior electrodes (with a rightward distribution in Kounios and Holcomb, 1994). Because processing these two types of nouns results in qualitatively different ERPs, the authors suggested that these data supported a modified version of Paivio's dual-coding model which accounted for the apparent effects of supportive vs. non-supportive sentence context.

West and Holcomb (2000) were able to tease apart the imageability portion of the concreteness N400 effect by using a manipulation of depth of processing: imagery, semantic, and surface. When participants were required to process using imagery (e.g., "It is easy to form a mental picture of an elephant/aptitude), the resulting ERPs to concrete (compared to abstract) nouns took the form of a greater negativity peaking between 650 and 750 ms, having a different scalp distribution than the traditional N400 (labeled an N700 component), which was taken to reflect a separate, more "imagistic" process being computed in this condition. The ERPs elicited by the semantic level of processing (e.g., "It is common for people to have an elephant/aptitude") conformed to the concreteness N400 effect seen in previous studies (Holcomb et al., 1999; Kounios & Holcomb, 1994), indicating to the authors

that the typical concreteness N400 effect is most likely the result of processing stimuli conceptually, thereby capturing the “concreteness” feature of the nouns used. The surface processing condition (e.g., “There is/is not an ‘n’ in the word elephant/aptitude”) failed to produce a difference between concrete and abstract nouns.

Thus, ERPs have proven appropriate to consistently index noun “concreteness” in English. Such concreteness N400 effects have also been found in Chinese (Zhang, Guo, Ding, & Wang, 2006), and Japanese (Nittono, Suehiro, & Hori, 2002), and possibly Spanish (Moreno & Kutas, 2005), among others.

Furthermore, ERPs have proven useful in examining the time-course of pronoun resolution using the lexical properties of (syntactic) gender (Osterhout & Mobley, 1995; Hagoort & Brown, 1999; Brown, van Berkum, & Hagoort, 1999; van Berkum, Brown, & Hagoort, 1999), animacy (Lamers, Jansma, Hammer, & Münte, 2006), and number (Barber & Carreiras, 2006). Finally, a recent study by Heine et al. (2006) demonstrates the validity of the use of ERPs to examine lexical frequency in pronominal coreference. They found that components elicited by pronouns coreferenced to less frequent antecedents resulted in greater N400s and P300s than waves in response to more frequent antecedents. The ERP methodology, thus, appears sensitive enough to permit a temporally precise examination of the timing of concreteness effects in pronoun coreference resolution.

2.1.6. The Current Study

The results of the studies described above suggest that a noun's concreteness may be a useful test for assessing the availability of an antecedent's conceptual information at a pronoun. ERPs constitute a novel approach for this testing, as there have been no ERP studies using concreteness to address pronoun resolution. Indeed, it seems probable that, if the concreteness N400 (cN400) can index concreteness on-line, it may also be capable of indexing concreteness effects during pronoun coreference.

It is not currently known whether an antecedent's concreteness is available to the parser immediately after encountering a co-referring pronoun, or whether there may be a delay in its availability. As mentioned above, there is some behavioral evidence that conceptual information from the antecedent may be available to the parser fairly early during pronoun processing. However, because of the temporal limits imposed by the behavioral methodologies – CMLP studies generally find evidence of priming based on RTs upwards of 600-700 ms following the presentation of the coreference – it is not clear that conceptual information is available immediately. Therefore, this study undertook to examine the following questions:

Question 1: Does the concreteness effect manifest at a pronoun during pronominal coreference?

Question 2: What is the time-course of pronominal concreteness effects?

Sentence pairs like the following were constructed to test these questions:

(7a) Phil knows all about **passion_a**. He loves to talk about **it_a**.

(7b) Phil knows all about **prison_c**. He loves to talk about **it_c**.

In (7a), “it” refers to “passion”, which is rated as being relatively abstract, but in (7b), “it” refers to “prison”, which is relatively concrete. Using this type of manipulation, if

the concreteness of the antecedent is “inherited” by the pronoun, it should be possible to visualize and compare the ERP signatures of the pronouns across the two conditions.

The following hypotheses were put forth:

Hypothesis 1: Pronouns coreferent with concrete nouns will “inherit” that concreteness; this will be manifested in a greater concreteness N400 effect to “concrete” pronouns than to “abstract”. On the other hand, if pronouns coreferenced with abstract antecedents are more difficult to resolve or integrate into the ongoing context, then a more typical N400 and/or a late positivity (P600/LPC) should be visible in that condition compared to ERPs from pronouns coreferenced with concrete nouns.

Hypothesis 2: If conceptual, semantic information is immediately available during pronoun resolution (i.e., bypassing the lexical activation stage; see Nicol and Swinney, 2003), then the concreteness of antecedents will be available immediately at the pronouns. Specifically, the concreteness N400 effect should be apparent within, or even prior to, the typical window of concreteness availability for nouns, but not later than the typical concreteness N400 window.

2.2. METHOD

This experiment was undertaken to elaborate the time-course of concreteness availability in pronoun coreference. In particular, because concreteness N400 effects have so far only been demonstrated in visually presented material, this experiment was

designed to test the hypothesis that the lexical property of concreteness is reactivated at a pronoun in the auditory modality.

Some effects in language processing are differentially elicited when processing visual versus auditory stimuli (e.g., Michael, Keller, Carpenter, & Just, 2001). For example, semantic priming effects begin earlier, are larger in size, and tend to last longer in the auditory modality than visual (Holcomb & Neville, 1990, 1991; Holcomb, Coffey, & Neville, 1992). Some effects, however, are not seen when participants process language in the auditory modality but only in the visual; this is particularly true in ERPs studies because components evoked by visual stimuli tend to be more sharply and consistently defined around an expected latency, while those evoked by equivalent auditory stimuli tend to be more broadly rounded, or “smeared” out, due to individual differences in time course of lexical recognition and processing. All the studies finding a concreteness N400 effect to language stimuli have employed the reading modality (Kounios & Holcomb, 1994; Holcomb et al., 1999; West & Holcomb, 2000; Zhang et al., 2006; Nittono, Suehiro, & Hori, 2002) with a single exception: van Schie, Wijers, Mars, Benjamins, and Stowe (2005) used single auditory word presentation to study concreteness effects on the retrieval of visual semantic information using ERPs. Finally, auditory sentence comprehension appears to be, at least to some degree, a more “natural”, fluent, and continuous process of comprehension (compared with, for example, word-by-word reading of visually presented sentences); it is therefore worth investigating whether effects seen during visual presentation are reflective of the same underlying processes as when presentation is auditory.

This experiment was therefore designed to investigate whether the concreteness effects seen in previous studies may be elicited during sentence processing in the auditory modality. Because these effects have never been demonstrated in the auditory modality, an attempt was made to maximize possible concreteness effects at the noun and pronoun in this study. This was accomplished in several ways.

First, stimuli were designed as minimal pairs, with an initial clause ending in either an abstract or concrete pronoun, and the second clause always ending with the pronoun “it”, as in (8) below:

(8) Phil knew all about prison_c/passion_a. He had personally experienced it_{c/a}.

This was expected to maximize effects, based on previous research in with so-called “sentence wrap-up effects” amplify structural processing strategies (Nicol, Swinney, Love, & Hald, 1997; Balogh, Zurif, Prather, Swinney, & Finkel, 1998).¹¹

Second, the sentences had no extraneous NPs, to avoid the possibility that other nouns might influence the concreteness effect at the critical word. Third, great care was taken to strategically place the pronoun “it” in a grammatically unambiguous position to ensure only one possible noun antecedent. For example, in an item such as, “John showed his fist. His friends were surprised by it”, the pronoun “it” might refer to the event instead of only the noun; “it” may plausibly refer to the fist or to John’s act of showing the fist. Similarly, cases in which “it” may be pleonastic, or semantically null (e.g., the “it” in “Sam knew it was about to start raining”) were

¹¹ It should also be noted that all 3 of the studies by Holcomb and colleagues (1994; 1999; 2000) that found a concreteness N400 effect in English did so with the target noun at the end of the sentence.

avoided. Finally, the task was chosen to maximize the probability of evoking the concreteness N400 effect by asking participants to actively rate imageability of each item. West and Holcomb (2000) demonstrated the greatest concreteness effect in conditions in which participants were asked to generate and assess mental images of the target words¹². Therefore, the task in the current experiment was intended enhance the likelihood of engaging semantically “deep” networks by asking participants to make imageability ratings to the items (see below).

2.2.1. Participants

In this study, twenty-five (25) UCSD students participated for course credit. Data from 4 participants were discarded due to an excessive number of movement or eye-blink artifact-influenced trials (> 30%); 6 participants’ data were discarded due to technical problems including experimenter error, faulty equipment, and data corruption. Consequently, data from fifteen (15) participants (5 male, 10 female) are included in the final analyses. Mean age of participants was 20.3 years (range 18-28). All were native speakers of English with no exposure to another language before age 6, and no visual, auditory, or neurological impairments. Two were left-handed¹³.

2.2.2. Stimuli

¹² In Experiment 2 of West et al., 2000, ERPs were collected as participants read and answered the sentence, “It is easy to come up with a mental image of...an elephant” vs. “...an element”.

¹³ Post hoc examination of their results indicated that their ERP patterns were not different than those of the right handed participants, so the data were collapsed together for the analyses reported herein.

Forty (40) “matched sets” of sentences (“items”) were created. A “matched set” conformed to the following pattern:

(9a) Bill knew all about passion(abs). He had personally experienced it(abs).

(9b) Bill knew all about prison(con). He had personally experienced it(con).

Matched sets were identical with the exception of the final word in the first sentence; that word was always one of the two “mate” target nouns (i.e., concrete vs. abstract). Mate nouns were matched based on length in syllables (all bisyllabic), stress pattern (6 abstract nouns were iambic; the remaining 34 abstract and 40 concrete nouns were trochees), and word-onset (30 out of the 40 mated noun pairs had identical onsets, while 10 pairs had non-identical onsets that shared at least one feature). Mated nouns were chosen to have extreme, opposing concreteness ratings¹⁴: mean concreteness rating for “concrete” nouns = 6.7; mean concreteness for “abstract” nouns = 3.3; this difference was significant ($t_{39} = -44.21, p < 0.0001$). Concrete and abstract nouns also differed in ratings of imageability (mean “concrete” = 6.3, mean “abstract” = 3.7); this difference was also significant ($t_{39} = -17.62, p < 0.0001$). Mates were matched for frequency (Francis & Kucera, 1982) such that no target noun differed from its mate in frequency more than 44 words per million (mean diff. = 1.33 per million); mean frequency for concrete and abstract nouns were 43 and 42 per million, respectively ($t_{39} = -0.566, p = 0.57$).

¹⁴ Concreteness ratings were obtained from a separate population of 22 UCSD undergraduates in an off-line, paper and pencil task in which the participants rated 152 nouns rated for concreteness using a 1 (very abstract) to 7 (very concrete) scale. Nouns with the most extreme ratings were selected for use in the ERP study.

As shown in (9) above, first sentences in the matched sets began with a proper noun to avoid concreteness effects resulting from unnecessary NPs. Second sentences were five or six words long ending in the pronoun “it”. To eliminate the possibility of “it” being temporarily ambiguous (e.g., as a possible pleonastic “it”), “it” in the second sentence was always the object of a transitive verb (e.g., “She tried to hide it”) or of a preposition (e.g., “He closely looked at it”).

Items were digitized on a computer using CoolEdit Pro v1.2 by a male, native English speaker with a standard American English accent. To enhance co-articulatory similarity between mated items, both sets of sentences in a matched set were recorded, and then the target noun for one item was digitally spliced into the file to replace its mate, with the end result being two items that were completely identical except for the target word. Target words did not differ in acoustic length ($\text{mean}_{\text{abstract}} = 573 \text{ ms}$, $\text{mean}_{\text{concrete}} = 544 \text{ ms}$, $t_{39} = 1.54$, $p = 0.13$). The first sentences in mated items had identical numbers of syllables (between 5 and 12 syllables, $\text{mean} = 8.53$) and did not differ with respect to rate of speech ($\text{mean}_{\text{abstract}} = 4.46 \text{ syll/sec}$; $\text{mean}_{\text{concrete}} = 4.54 \text{ syll/sec}$; $t_{39} = -1.56$, $p = 0.13$). Within each item, the last word in the first sentence (i.e., the target noun) was followed by a 1500 ms pause, followed by the second sentence. A jittered window of silence (1250 ms, 1500 ms, or 1750 ms) followed the end of the second sentence (i.e., the offset of the pronoun “it”), followed by a beep as a prompt for participant responses (see below).

Items were counterbalanced across both halves of the experiment in order to maximize the number of intervening items between mates. Ten practice sentences were presented at the very beginning of the study to allow participants to become

familiar with the task. Eighty filler stimuli were made up of sentence pairs having a variety of structures to reduce pattern learning and strategic processing in participants.

2.2.3. Procedure

Participants were seated comfortably in a sound-proof booth approximately 1.65 meters from a pair of stereo speakers. Participants were instructed to listen to the stimuli and make a binary, button-press judgment after each item depending on whether the sentences represented events for which they found it “easy” to generate an immediate mental image versus “not as easy”. Two examples were given to ensure that the participants understood that the task was not to force images for all sentence pairs but to judge whether an image easily came to mind for each item.¹⁵ Participants were asked to remain still without blinking while listening to the entire item and press the button only after they heard the beep. After 10 practice items to acquaint participants with the task, the speakers presented 160 sentence pairs during 3 blocks: 10 practice items, and then 80 experimental (40 containing a concrete noun and 40 containing an abstract noun), counterbalanced and pseudo-randomized among the 80 fillers, such that no experimental condition was repeated more than two times in a row. Each block began with 3 filler sentences to ensure that participants were fully engaged in the task before they heard experimental items. Response hand for “very imageable” versus “not imageable” was counterbalanced across participants.

¹⁵ The two example items were:

Concrete: “The gymnast polished her medal. She was very proud of it.”

Abstract: “The man talked about memory. He was clearly an expert on it.”

2.2.4. ERP procedure

An electrode cap (QuikCap, CompuMedics, Inc.) embedded with 32 sintered Ag/AgCl electrodes was applied to the scalp, with all electrodes referenced to the left mastoid. The sites involved a modified International 10-20 System, including midline (Fz, FCz, Cz, CPz, Pz, and Oz) and lateral sites (FP1/2, F3/4, F7/8, C3/4, T7/8, P3/4, P7/8, and O1/2), as well as at FC3/4, FT7/8, CP3/4, and TP7/8. In addition, bipolar loose electrodes were applied to the outer canthus of the both eyes and above and below the left eye to monitor vertical and horizontal eye movements, respectively.

The EEG signal was amplified 100 X with a SynAmps amplifier (3 dB cutoff, bandpass filtered on-line between 0.15 and 30 Hz), digitized on-line at 500 Hz. Average ERPs were formed off-line from trials free of EOG and movement artifact ($>\pm 100\mu\text{V}$). As mentioned in “Participants” above, 4 participants exhibited excessive amounts of artifact (~50% of epochs were unusable due to movement and/or eye blinks) and their data were excluded from later analysis. The remaining participants showed a low rate of lost trials due to artifacts ($<10\%$ across conditions).

2.2.5. Data analysis

The ERP data to the conditions of interest (concrete and abstract nouns and pronouns) were quantified by measuring the mean amplitude (in μV) within three latency windows that were determined, based on a significant literature (see, e.g., Friederici, 2002), to be most appropriate for investigating the concreteness effect: 100-300 ms, 300-500 ms, and 500-1000 ms. Separate grand average analyses were performed for each time window separately for nouns and pronouns. Two-way

repeated measures analysis of variance (ANOVA) was used for each initial analysis, using concreteness (Concrete vs. Abstract) and Grouping variables as within-subjects factors. The within-subject Grouping variables were: laterality (Central vs. Left vs. Right), anteriority (Anterior: F3/4, F7/8, FC3/4, FT7/8, Fz, FCz; Central: Cz, C3/4, CPz, CP3/4; and Posterior: T7/8, TP7/8, Pz, P3/4, P7/8, Oz, O1/2), and scalp region (Left Frontal: F3, F7, FC3, FT7; Right Frontal: F4, F8, FC4, FT8; Left Posterior: P3, P7, T7, TP7; Right Posterior: P4, P8, T8, TP8; and Centroparietal: Pz, CPz, Cz, C3/4, CP3/4); the region grouping was chosen to maximize the ability to visualize (E)LAN, N400, and P600 effects. The Huynh-Feldt (HF) correction was imposed whenever the number of degrees of freedom in the numerator was greater than 1; all reported *p* values are corrected.

2.3. RESULTS

2.3.1. Behavioral results

Attempting to image the items was the only task demand, as this was considered a sufficient for generating the expected concreteness effects (cf. West & Holcomb, 2000). Therefore, percent correct responses were not analyzed.

2.3.2. ERP results

Results reported here are based on analyses of the unfiltered data unless otherwise specified. Scalp distributions of ERP results for nouns are shown in Figure 3.

2.3.2.1. NOUNS, 100-300 ms

Two-factor repeated measures ANOVA (rmANOVA) with concreteness X laterality, concreteness X anteriority, and concreteness X region, all as within-subjects measures, showed no significant main effects or interactions in this window.

2.3.2.2. NOUNS, 300-500 ms

Two-factor rmANOVA with concreteness X laterality as within-subjects factors found no main effects and no interactions were reached.

Two-factor rmANOVA with concreteness X anteriority found no main effects or interactions.

A two-factor rmANOVA with concreteness X region found no main effects or interactions.

2.3.2.3. NOUNS, 500-1000 ms

A two-factor rmANOVA with concreteness X region found a main effect of region [$F(4, 56) = 7.117, p = 0.0168$]. This manifested as a difference between the Left Anterior and Centroparietal groups, with mean amplitudes being least positive at Left Anterior electrodes and most positive at Centroparietal electrodes (mean amplitude = $0.237 \mu V$ vs. $2.451 \mu V$, respectively). There was also a difference between Right Anterior and Centroparietal groups (mean amplitude = $0.780 \mu V$ vs. $2.451 \mu V$, respectively). Left Posterior ($1.563 \mu V$) and Right Posterior ($1.605 \mu V$) electrode groups were non-different from other regions. No interactions were found.

2.3.2.4. PRONOUNS, 100-300 ms

Scalp distributions of ERP results for pronouns are shown in Figure 4.

Two-factor rmANOVA with concreteness and laterality found a main effect of laterality [$F(1, 14) = 9.353, p = 0.0085$], with Left hemisphere electrodes being significantly more negative than Right ($-0.114 \mu\text{V}$ vs. $0.411 \mu\text{V}$).

Similarly, two-factor rmANOVA with concreteness and anteriority found a main effect of anteriority [$F(2, 28) = 25.279, p = 0.0002$], revealing an overall greater negativity at Anterior electrodes ($-0.478 \mu\text{V}$) compared to both Central ($1.195 \mu\text{V}$) and Posterior ($0.652 \mu\text{V}$) electrodes.

Finally, two-factor rmANOVA with concreteness and region as within-subjects factors produced a main effect of region in this time window [$F(4, 56) = 21.581, p = 0.0003$], with the Left Frontal region being significantly more negative ($-1.005 \mu\text{V}$) than Right Frontal ($-0.146 \mu\text{V}$), Left Posterior ($0.034 \mu\text{V}$), Right Posterior ($0.242 \mu\text{V}$), and Centroparietal ($1.282 \mu\text{V}$) electrodes. Furthermore, Centroparietal electrodes were significantly more positive than Left Frontal, Left Posterior, Right Frontal, and Right Posterior. No interactions were noted. Significantly, there were no main effects found for concreteness (see Discussion) in this analysis.

2.3.2.5. PRONOUNS, 300-500 ms

rmANOVA with concreteness and laterality as within-subjects factors found a borderline main effect of concreteness [$F(1, 14) = 3.258, p = 0.0926$], with Concrete Pronouns overall more negative ($-0.402 \mu\text{V}$) than Abstract Pronouns ($0.346 \mu\text{V}$). A main effect of laterality was also found [$F(1, 14) = 8.315, p = 0.0120$], with Left

hemisphere electrodes having greater negative mean amplitude ($-0.381 \mu\text{V}$) than Right hemisphere electrodes ($0.325 \mu\text{V}$).

The rmANOVA with concreteness and anteriority as factors found a borderline main effect of concreteness [$F(1, 14) = 4.183, p = 0.0601$], with Concrete Pronouns ($-0.246 \mu\text{V}$) being overall more negative than Abstract ($0.644 \mu\text{V}$). The factor of anteriority also showed a significant main effect [$F(2, 28) = 17.333, p = 0.0010$], with Anterior electrodes ($-1.130 \mu\text{V}$) more negative than Central ($0.870 \mu\text{V}$) and Posterior ($0.856 \mu\text{V}$) electrodes, overall. A borderline interaction obtained between concreteness and anteriority [$F(2, 28) = 3.501, p = 0.0824$]. The interaction manifested itself as a greater negative amplitude for Concrete ($-1.774 \mu\text{V}$) vs. Abstract ($-0.486 \mu\text{V}$) Pronouns (concreteness effect) at Anterior electrodes [$F(1, 28) = 3.097, p = 0.0894$], but no significant differences in amplitude were demonstrated by Concrete and Abstract Pronouns at Posterior ($0.658 \mu\text{V}$ vs. $1.053 \mu\text{V}$) or Central ($0.376 \mu\text{V}$ and $1.364 \mu\text{V}$, respectively) electrodes.

Finally, two-factor rmANOVA with concreteness and region showed a main effect of region [$F(4, 56) = 12.539, p = 0.0027$], with Left Anterior electrodes being significantly more negative ($-1.510 \mu\text{V}$) than Left Posterior ($-0.101 \mu\text{V}$), Right Posterior ($0.394 \mu\text{V}$), and Centroparietal ($1.020 \mu\text{V}$) regions. Right Anterior ($-0.832 \mu\text{V}$) electrodes were also significantly more negative than Centroparietal, but were not different than Left Anterior electrodes. Also, a borderline effect of concreteness [$F(1, 14) = 3.982, p = 0.0658$] demonstrated that Concrete Pronouns were overall more negative ($-0.661 \mu\text{V}$) than Abstract Pronouns ($0.250 \mu\text{V}$). Although Concrete Pronouns were more negative than Abstract Pronouns *overall*, ANOVAs of

concreteness at each region approached significance only at the Right Anterior group of electrodes [$F(1, 28) = 3.734, p = 0.0635$].

2.3.2.6. PRONOUNS, 500-1000 ms

In the two-factor rmANOVA with concreteness and laterality as within-subjects factors, an effect of laterality was found [$F(1, 14) = 14.067, p = 0.0022$], as was an interaction between concreteness and laterality [$F(1, 14) = 10.269, p = 0.0064$]. Overall ERPs over the Left hemisphere ($-0.251 \mu V$) were more negative than over the Right ($0.670 \mu V$). Furthermore, Concrete Pronouns ($0.100 \mu V$) were significantly less positive than Abstract Pronouns ($1.239 \mu V$) over the Right hemisphere, and the amplitudes of ERPs to Concrete ($-0.409 \mu V$) and Abstract ($-0.094 \mu V$) Pronouns were not significantly different in the Left hemisphere. A borderline effect of concreteness [$F(1, 14) = 3.248, p = 0.0931$] hinted that Concrete Pronouns were overall more negative ($-0.154 \mu V$) than Abstract Pronouns ($0.537 \mu V$).

Two-factor rmANOVA with concreteness and anteriority found a main effect of concreteness [$F(1, 14) = 4.528, p = 0.0516$], showing that Concrete ($-0.050 \mu V$) Pronouns were overall more negative than Abstract ($0.832 \mu V$). A main effect of anteriority also emerged [$F(2, 28) = 6.283, p = 0.0251$], with Anterior ($-0.378 \mu V$) electrodes exhibiting significantly more negativity than Central ($0.882 \mu V$) or Posterior ($0.669 \mu V$) electrodes.

Finally, two-factor rmANOVA with concreteness and region as within-subjects factors demonstrated borderline significance for concreteness [$F(1, 14) = 3.447, p = 0.0845$], and a main effect for region [$F(4, 56) = 7.067, p = 0.0172$], with Left

Anterior ($-0.807 \mu\text{V}$) and Left Posterior ($-0.279 \mu\text{V}$) clusters being significantly more negative than both Centroparietal ($0.966 \mu\text{V}$) and Right Posterior ($0.731 \mu\text{V}$) electrode groups. The interaction between concreteness and region approached significance [$F(4, 56) = 4.026, p = 0.0620$], with the Right Anterior region showing a drop in amplitude considerably greater when comparing Abstract to Concrete (0.812 to $-0.791 \mu\text{V}$, $\text{diff.} = 1.603$) than for any other region (Centroparietal: 1.456 vs. $0.457 \mu\text{V}$, $\text{diff.} = 0.999$; Left Frontal: -0.676 vs. $-0.938 \mu\text{V}$, $\text{diff.} = 0.262$; Left Posterior: -0.085 vs. $-0.474 \mu\text{V}$, $\text{diff.} = 0.559$; Right Posterior: 1.251 vs. $0.212 \mu\text{V}$, $\text{diff.} = 1.039$). As in the previous window, this apparent “concreteness effect” was significant only in the Right Anterior electrode group [$F(1, 28) = 4.095, p = 0.0527$].

Because this study is also concerned with the issue of *when* concreteness information is available to the parser, post-hoc analyses were conducted, to narrow down the temporal window in which the difference in concreteness in the pronoun condition becomes apparent. The concreteness N400 effect in pronouns reached significance in the 300-500 ms window, so it is reasonable to infer that concreteness information is available to the parser by this point in time, at the latest. Post-hoc rmANOVA in the 300-400 ms window found a trend toward significance for concreteness [$F(1, 14) = 2.760, p = 0.1189$], with concrete pronouns significantly more negative ($-0.170 \mu\text{V}$) than abstract pronouns ($0.630 \mu\text{V}$) in this time window. It thus appears that the cognitive process underlying the pronoun concreteness effect is not apparent in the 200-300 ms window but does become apparent, as indexed by a trend beginning approximately 300-400 ms after the onset of the pronoun.

2.4 DISCUSSION

This experiment employed sentence processing in the auditory modality to test the time course of concreteness “reactivation” in pronouns coreferenced to antecedents with widely different concreteness ratings. A notable result is that the stimuli did not elicit a concreteness N400 ERP effect to Nouns in any of the time windows examined. Perhaps most surprisingly, a concreteness N400 effect was greatest over centroparietal electrodes rather than over anterior electrodes, as has consistently been found by Holcomb and colleagues and others (Holcomb et al., 1999; West and Holcomb, 2000; Nittono et al., 2002). This scalp distribution is less consistent with the cN400 but more consistent with the typical N400 effect (see Figure 7).

Across both abstract and concrete sentences, nouns produced greater positivity over centroparietal scalp regions and greater negativity frontally. The lack of a robust concreteness N400 effect between concrete and abstract nouns was noteworthy, for two reasons: because the concreteness N400 has been produced in similar experimental designs, though admittedly never in the auditory modality¹⁶; and because the target words were chosen to have widely divergent concreteness ratings. Furthermore, it was expected that the task of requiring participants to actively attempt to visualize the sentences would at the very least engage the imagistic system described by West and Holcomb (2000) sufficiently to produce some differences in the ERP trace. It is possible that the overall negativity toward anterior electrodes

¹⁶ In fact, a concreteness effect was produced by similar stimuli in a pilot study conducted in the visual modality.

reflects this process, but without specific conditions to test that possibility, it remains unclear.

Nevertheless, that a concreteness N400 did not appear in this study is, albeit unexpected, not entirely without explanation. Previous studies supporting the context-availability model (Kieras, 1978) have suggested that when concrete and abstract nouns are equally contextually supported, the concreteness effect disappears (Schwanenflugel, 1991; Schwanenflugel & Stowe, 1989). This outcome is presumably because when they are placed in an equally permissible environment, concrete words – which are claimed to be less ambiguous and more accessible than abstract words in the majority of contexts – lose that relative availability advantage over abstract words. Thus, if the sentence context is equally unbiased toward words of either concreteness, then word concreteness should not have an impact on lexical activation processes. Indeed, this possibility is supported by the results of Holcomb et al., (1999), who found concreteness N400 effects only in sentences for which the final (abstract or concrete) word was anomalous; Holcomb et al. state that “a supportive sentence context wiped out all ERP evidence of concreteness effects (p.730).” In an attempt to make both concrete and abstract words equally unpredictable, the sentences in the current study were deliberately and carefully chosen so as to have equally biasing context to support either concrete or abstract words. Coupled with the possible attenuating influence of the auditory modality, this seems to be a likely explanation for the lack of concreteness effects in the current study. Once a semantic context is established by the noun, however, the pronoun appears to be sensitive to the concreteness manipulation.

The results at the pronoun are rather compelling. Within pronouns, a broadly distributed (but greater anteriorly, and to the right) negativity was found in response to pronouns following concrete antecedents compared to deflections produced by pronouns with abstract antecedents. There was some indication of less positivity for concrete pronouns compared to abstract pronouns in the 100-300 ms window, though the effect was marginal; post-hoc tests revealed that this failed to reach significance even in the 200-300 ms window. A concreteness effect became significant in the 300-500 ms window, and continued to be significant in the 500-1000 ms window (and, on visual inspection, throughout the 1500 ms epoch). This effect was maximal anterior and to the right, with the effect approaching or reaching significance at the Right Anterior electrode group in both time windows. This distribution is consistent with the concreteness N400 elaborated by Holcomb and colleagues (Holcomb et al., 1999; West & Holcomb, 2000). Finding what appears to be a concreteness N400 effect in the ERPs to pronouns following concrete vs. abstract antecedents is a noteworthy result, not only because it supports the initial hypothesis that concreteness can be indexed at a co-referring pronoun. This appears to be the first time in the psycholinguistic literature that an antecedent's concreteness has been shown to be indexed at a downstream pronoun.

In the pronoun condition, the post-hoc ANOVA revealed that the cN400 effect approached significance in the 300-400 ms window and not sooner. This result is notable, because it demonstrates that the concreteness N400 effect in pronouns is not an illusory effect caused by a prolonged, stimulus-long negative shift in the ERP following the concrete antecedent (compared to the abstract one). This result also

argues that conceptual information, such as concreteness, is available to the parser almost immediately following the processing of a pronoun. Given that the time course for a noun concreteness effect in the Holcomb et al. studies tends to be 400-700 or - 800 ms following visual onset of the target noun, and because the conceptual information begins to be available, at the latest, ~300 ms after a pronoun is encountered in this study (i.e., perhaps 100 ms earlier at the pronoun), it is possible that antecedent reactivation at a pronoun may not proceed according to the same time course as typical lexical (noun) activation. Unfortunately, the current study was unable to elicit a cN400 in response to the nouns, so a direct comparison is impossible. Although only conjecture at this point, the timing of the pronoun cN400 hints that, whereas lexical activation of an antecedent requires early stages of word form recognition (including neighborhood effects, etc.) before conceptual, semantic information becomes available (by ~400-800 ms), conceptual information coming from antecedent reactivation at a pronoun may be available earlier, by ~300-400 ms. This is consistent with results from Nicol and Swinney (1989) and Love and Swinney (1995), who suggest that antecedent reactivation is not simply lexical reactivation but “skips” the usual early form-driven steps and accesses a deeper, conceptual representation. Taken as one piece of a larger whole, these data appear to support models that consider antecedent reactivation qualitatively different from lexical activation (cf. Simner & Smyth, 1999).

These results strongly suggest that concreteness can be considered on a par with other lexical and conceptual properties of antecedents that are reactivated, or available for processing, at coreferent pronouns. Although these data demonstrate that

concreteness information is *available* to be used by the parser in coreference assignment in a timely manner, the claim here is certainly not that concreteness itself is, *de facto*, obligatorily considered by the parser to drive coreference computation. Concreteness is viewed only as in index of the semantic properties which are available. Indeed, it seems unlikely that there are many natural situations in which antecedent concreteness would be the key factor in determining which antecedent is selected from the candidates. Nevertheless, future designs using concreteness to index pronominal coreference in sentences with more than one plausible antecedent can easily be imagined, using concreteness differences in antecedents to determine which antecedent assignment is “preferred”. That is, if a noun’s concreteness is obligatorily and immediately reconstituted at the pronoun, then ERPs to pronouns in sentences with both concrete and abstract antecedents in specified structural positions may allow the visualization of which antecedent is initially perceived to be appropriate.

These data suggest that there is a neural processing distinction between pronouns coreferenced to concrete vs. abstract antecedents, and they indicate that this computation is performed in a circumscribed period of time. These results motivate the need to investigate behavioral correlates for concreteness effects in pronominal coreference, as no such studies can be found in the literature. Another line of buttressing evidence would come from reaction time studies to determine whether the concreteness effect does, in fact, influence on-line processing using a more traditional methodology. Because of its previous successes in this area, a likely candidate methodology is Cross Modal Lexical Priming.

In sum, a substantial literature supports the computation of lexical properties such as gender, number, animacy, and lexical frequency information during on-line pronoun resolution in sentence processing. This study represents a first step in verifying that the conceptual property of concreteness in antecedents is also available for those computations immediately at the pronoun.

Table 2.1: Attributes of concrete and abstract nouns

	Mean Concreteness Rating (1 – 7)	Mean Frequency (per million)	Mean Duration (in ms)	Mean # syll/sec
Concrete	6.7	43	544	4.54
Abstract	3.3	42	573	4.46

ms = milliseconds; syll = syllables; sec = second

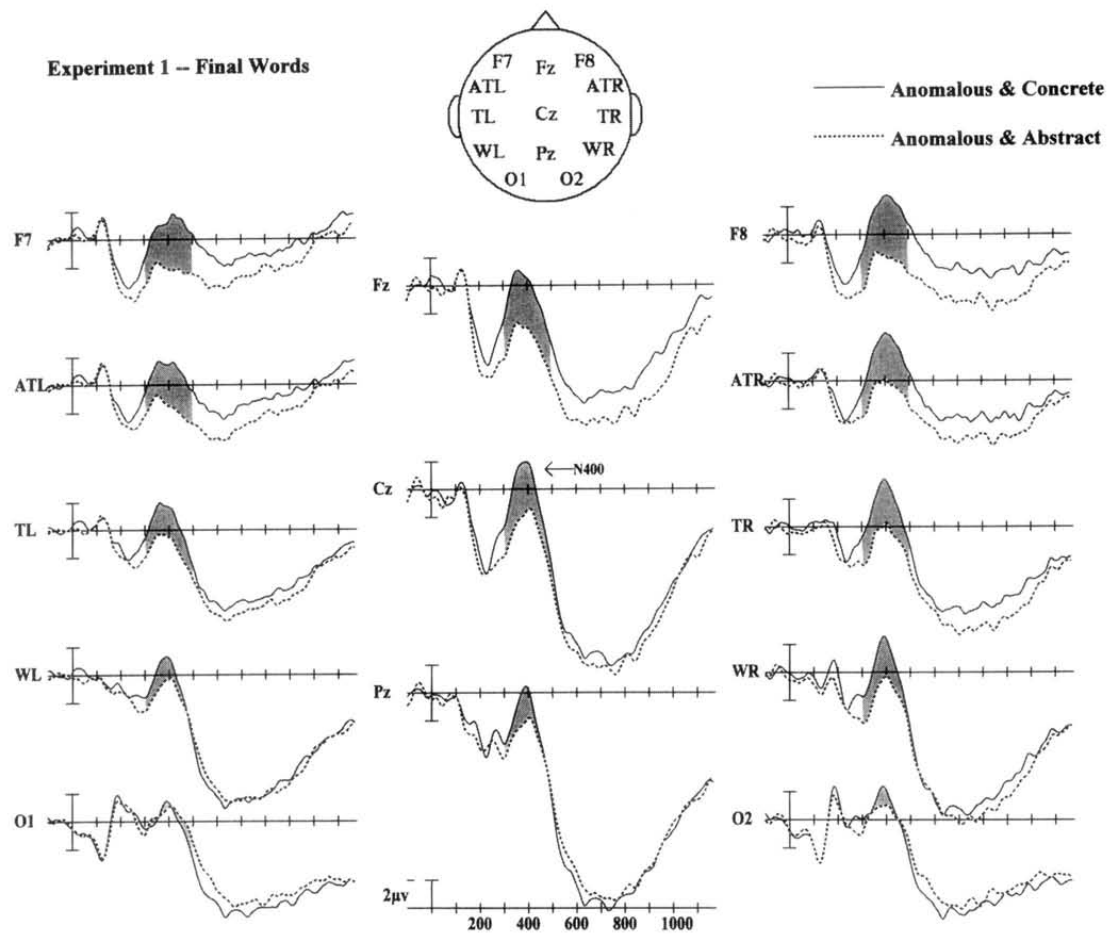


Figure 2.1: The Concreteness N400 effect; figure taken from Holcomb et al. (1999). Note that, in the Anomalous condition, sentences ending in Concrete nouns resulted in significantly larger negativities in the N400 time window (300-500 ms) than Abstract nouns.

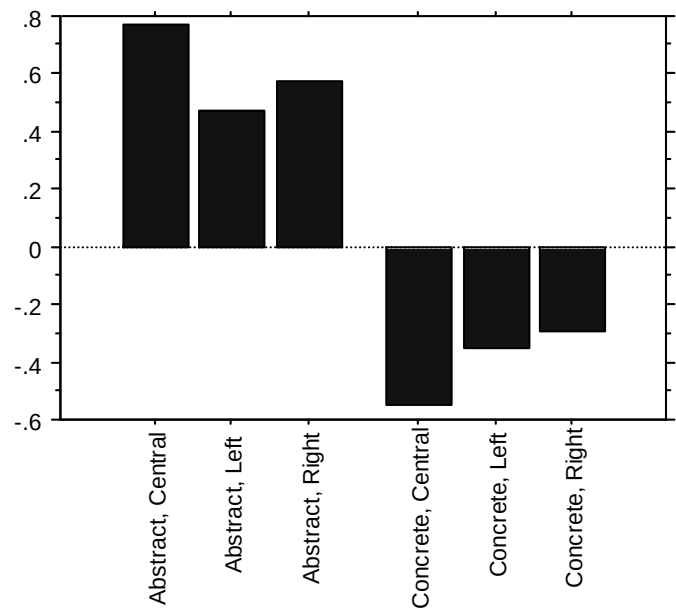


Figure 2.2: Mean amplitude for Abstract (left) and Concrete Nouns (right) over across both hemispheres in the 300-500 ms window. Amplitude units are in microvolts (μV).

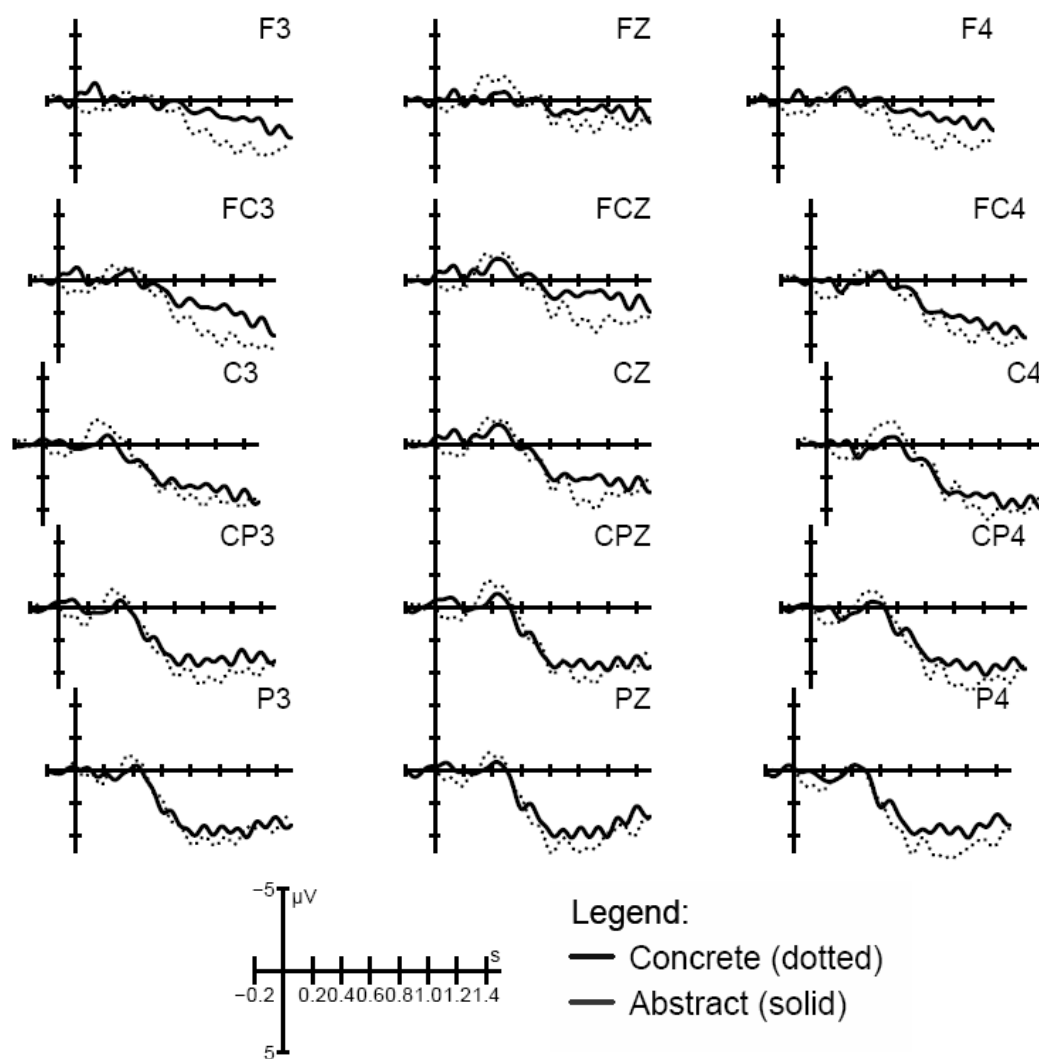


Figure 2.3: Grand average ERP waveforms (referenced to linked mastoids) elicited by antecedent nouns rated as being strongly “concrete” or “abstract” in auditorily presented sentences. For visualization purposes, these waveforms have been bandpass filtered at 0.05 – 9 Hz. ERPs to concrete nouns are represented in blue; ERPs to abstract nouns are red.

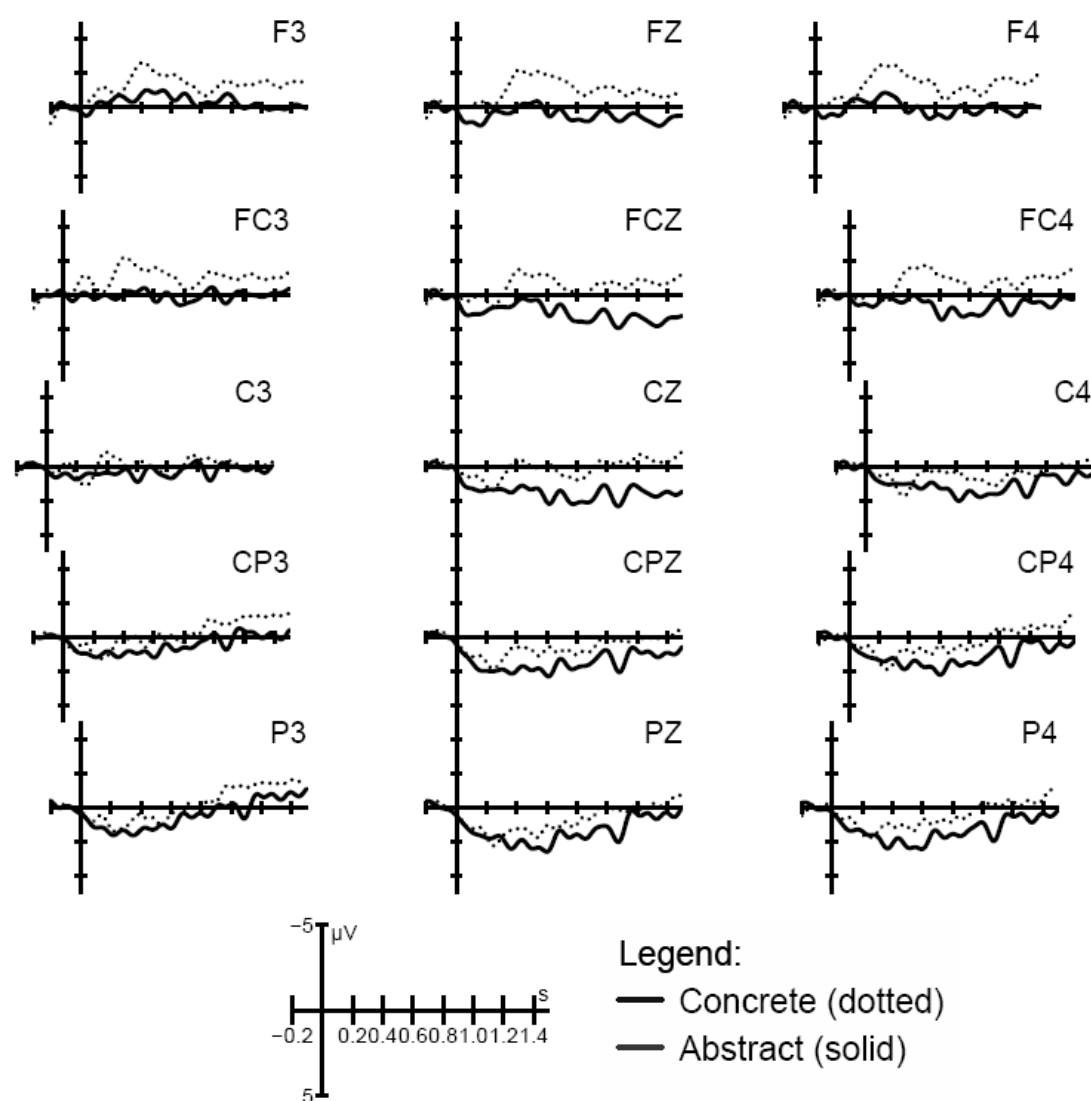


Figure 2.4: Grand average ERP waveforms (referenced to linked mastoids) elicited by pronouns co-referent with antecedent nouns rated as being strongly “concrete” or “abstract” in auditorily presented sentences. For visualization purposes, these waveforms have been bandpass filtered at 0.05 – 9 Hz. ERPs to “concrete” pronouns are represented in blue; ERPs to “abstract” pronouns are red.

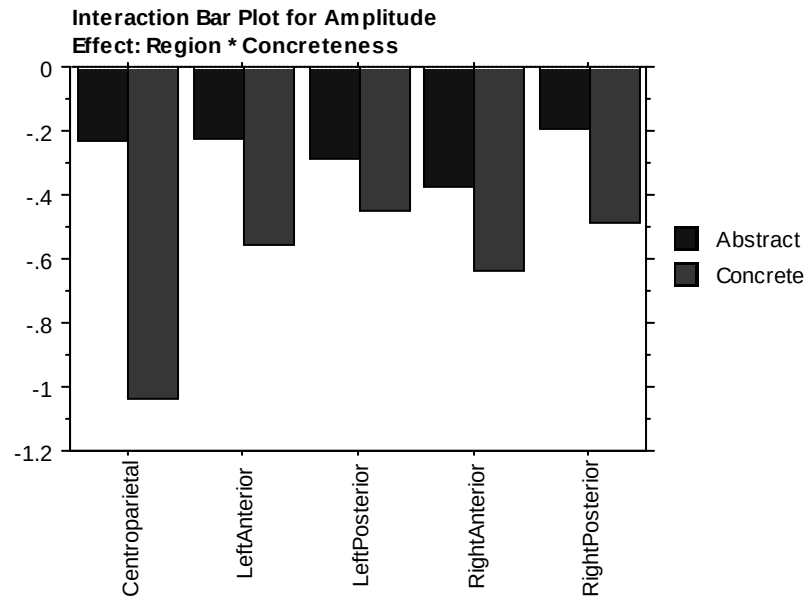


Figure 2.5: Mean amplitude for Abstract Pronouns (black) and Concrete Pronouns (gray) over 5 scalp Regions (Centroparietal, Left Anterior, Left Posterior, Right Anterior, and Right Posterior) in the 300-500 ms window. Amplitude is in microvolts (μV).

APPENDIX A

EXPERIMENTAL MATERIALS

Paul often talked about the affair. He could not forget it.
Paul often talked about the ocean. He could not forget it.

Julie knows everything about the ancient compound. She tells everyone about it.
Julie knows everything about the ancient compass. She tells everyone about it.

Eric had a crooked feature. He was embarrassed by it.
Eric had a crooked finger. He was embarrassed by it.

Alycia lied about her habit. She wanted to hide it.
Alycia lied about her candy. She wanted to hide it.

Chris had a secret pleasure. He tried to hide it.
Chris had a secret pocket. He tried to hide it.

Jacob remembered the tribute. He couldn't describe it.
Jacob remembered the pistol. He couldn't describe it.

Fred was proud of his virtue. He bragged about it.
Fred was proud of his baby. He bragged about it.

Sarah has a temper. She tried to hide it.
Sarah has a ticket. She tried to hide it.

Eric proudly displayed his talent. He's very proud of it.
Eric proudly displayed his theater. He's very proud of it.

Bill knew all about passion. He had experienced it.
Bill knew all about prison. He had experienced it.

Ken watched the mischief. He wanted to stop it.
Ken watched the monkey. He wanted to stop it.

Tanya studied the antique image. She found something in it.
Tanya studied the antique window. She found something in it.

Brittany didn't have the patience. She had never needed it.
Brittany didn't have the pitcher. She had never needed it.

Jim was stunned by the judgment. He had never expected it.
Jim was stunned by the garden. He had never expected it.

Felix knew of the secret motive. He tried to hide it.
Felix knew of the secret mirror. He tried to hide it.

Agatha received a magical blessing. She often brags about it.
Agatha received a magical berry. She often brags about it.

Colin finally emerged from his despair. He gladly left it.

Colin finally emerged from his doorway. He gladly left it.

Ann likes to show her courage. She often brags about it.
Ann likes to show her cabin. She often brags about it.

Michael was excited about the new prospect. He had constant thoughts of it.
Michael was excited about the new penny. He had constant thoughts of it.

Andy couldn't hide his fury. You could not ignore it.
Andy couldn't hide his forehead. You could not ignore it.

Bob closely studied the old custom. He learned everything about it.
Bob closely studied the old carpet. He learned everything about it.

Stan found a hidden fortune. He closely looked at it.
Stan found a hidden football. He closely looked at it.

Alan aligned the array. He visually inspected it.
Alan aligned the arrow. He visually inspected it.

Jenny had a strange belief. She tried not to reveal it.
Jenny had a strange bedroom. She tried not to reveal it.

Leslie was almost out of credit. She eventually ran out of it.
Leslie was almost out of coffee. She eventually ran out of it.

Phil clearly understood the notion. He had deep knowledge about it.
Phil clearly understood the motor. He had deep knowledge about it.

Sally remembered the strange feeling. She wrote all about it.
Sally remembered the strange table. She wrote all about it.

Tracy studied the hidden factor. She rarely talks about it.
Tracy studied the hidden forest. She rarely talks about it.

Mick thought about his last duty. He could not forget it.
Mick thought about his last dollar. He could not forget it.

Patricia didn't expect such a pursuit. She tried to manage it.
Patricia didn't expect such a package. She tried to manage it.

Karen was short on beauty. She wanted more of it.
Karen was short on butter. She wanted more of it.

Mallory smiled at the neighbor's kindness. She commented on it.
Mallory smiled at the neighbor's kitten. She commented on it.

Johnny tried to hide his obvious sorrow. He was embarrassed by it.
Johnny tried to hide his obvious stomach. He was embarrassed by it.

George researched the old legend. He wrote often about it.
George researched the old lumber. He wrote often about it.

Ralph was sick of the silence. He got rid of it.

Ralph was sick of the sandwich. He got rid of it.

Carol was concerned about the excess. She complained loudly about it.
Carol was concerned about the engine. She complained loudly about it.

Julio owned a very small amount. He wasn't proud of it.
Julio owned a very small island. He wasn't proud of it.

Amber avoided the danger. She moved away from it.
Amber avoided the kitchen. She moved away from it.

Alan looked for the bargain. He finally found it.
Alan looked for the hammer. He finally found it.

Marty really wanted more wisdom. He eventually did get it.
Marty really wanted more liquor. He eventually did get it.

REFERENCES

- Arnold, JE, Eisenband, JG, Brown-Schmidt, S, & Trueswell, JC (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13-B26.
- Balogh, J, Zurif, E, Prather, P, Swinney, D, & Finkel, L (1998). Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language*, 61, 169-182.
- Carreiras, M, Garnham, A, Oakhill, J, & Cain, K (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, 49A, 639-663.
- Carreiras, M, Gernsbacher, MA, & Villa, V (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review*, 2, 124-129.
- Clark, HH & Begun, JS (1971). The semantics of sentence subjects. *Language and Speech*, 14, 34-46.
- De Vincenzi, M (1999). Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research*, 28, 537-555.
- Garnham, A (2001). Mental models and the interpretation of anaphora. Hove: Psychology Press.
- Garnham, A, Oakhill, J, & Cruttenden, H (1992). The role of implicit causality and gender cue in the interpretation of pronouns. *Language and Cognitive Processes*, 73, 231-255.
- Garnham, A, Oakhill, J, Ehrlich, M-F, & Carreiras, M (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, 34, 41-62.
- Gernsbacher, MA, & Hargreaves, DJ (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27, 699-717.
- Gernsbacher, MA, Hargreaves, DJ, & Beeman, M (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, 28, 735-755.

- Greene, SB, McKoon, G & Ratcliff, R (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 182, 266-283.
- Hagoort, P, & Brown, CM (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 715-728.
- Heine, A, Tamm, S, Hofmann, M, Hutzler, F, & Jacobs, AM (2006). Does the frequency of the antecedent noun affect the resolution of pronominal anaphors? An ERP study. *Neuroscience Letters*, 400, 7-12.
- Holcomb, PJ, Kounios, J, Anderson, JE, & West, WC (1999). Dual-coding, context-availability, and concreteness effect in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 721-742.
- Kennison, SM & Trofe, JL (2003). Comprehending pronouns: a role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, 32, 355-378.
- Kieras, D (1978). Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin*, 85, 532-554.
- Kounios, J & Holcomb, PJ (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 804-823.
- Kroll, JF, & Merves, JS (1986). Lexical access for concrete and abstract nouns. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 92-107.
- Kucera, H, & Francis, WN (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press
- Kutas, M, Van Petten, C, & Kluender, R (2006). Psycholinguistics Electrified II (1994-2005). In MA Gernsbacher & M Traxler (Eds.), Handbook of Psycholinguistics, 2nd Edition. New York: Elsevier Press.
- Lamers, MJA, Jansma, BM, Hammer, A, & Münte, TF (2006). Neural correlates of semantic and syntactic processes in the comprehension of case marked pronouns: Evidence from German and Dutch. *BMC Neuroscience*, 7: 23.

- Love, T, & Swinney, D (1996). Coreference processing and levels of analysis in object-relative constructions; Demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, 25, 5-24.
- Michael, EB, Keller, TA, Carpenter, PA, & Just, MA (2001). fMRI investigation of sentence comprehension by eye and ear: Modality fingerprints on cognitive processes. *Human Brain Mapping*, 13, 239-252.
- Nicol, J, & Swinney, D (1989). The role of structure and coreference assignment during sentence comprehension. Sentence Processing [Special Issue]. *Journal of Psycholinguistic Research*, 18, 5-19.
- Paller, K (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differently by directed forgetting. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 1021-1032.
- Paller, K.A., & Kutas, M. (1992). Brain potentials during retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 4, 375-391.
- Pynte, J & Colonna, S (2000). Decoupling syntactic parsing from visual inspection: the case of relative clause attachment in French. In A Kennedy, R Radach, D Heller, & J Pynte (Eds.), Reading as a perceptual process (pp. 529-547). Oxford: Elsevier.
- Reynolds, DJ, Garnham, A, & Oakhill, J (2006). Evidence of immediate activation of gender information from a social role name. *The Quarterly Journal of Experimental Psychology*, 59, 886-903.
- Richardson, JTE (1976). The effects of stimulus attributes on the latency of word recognition. *British Journal of Psychology*, 67, 315-325.
- Samson, D & Pillon, A (2004). Orthographic neighborhood and concreteness effects in the lexical decision task. *Brain and Language*, 91, 252-264.
- Shillcock, RC (1982). The on-line resolution of pronominal anaphora. *Language and Speech*, 25, 385-401.
- Simner, J & Smyth, R (1999). Phonological activation in Anaphoric Lexical Access (ALA). *Brain and Language*, 68, 40-45.
- Smyth, R (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23, 197-229.

- Sturt, P (2003). A new look at the syntax-discourse interface: the use of binding principles in sentence processing. *Journal of Psycholinguistic Research*, 32, 125-139.
- Tyler, LK, Voice, JK, & Moss, HE (2000). The interaction of meaning and sound in spoken word recognition. *Psychological Bulletin and Review*, 7, 320–326.
- Van Berkum, JJA, Brown, CM, & Hagoort, P (1999). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research*, 28, 555-571.
- van Gompel, RGP & Majid, A (2004). Antecedent frequency effects during the processing of pronouns. *Cognition*, 90, 255-264.
- van Gompel, RL & Liversedge, SP (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 128-139.
- Walker, I, & Hulme, C (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1256–1271.
- West, WC & Holcomb, PJ (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience*, 12, 1024-1037.

Chapter 3

Abstract

This study examines the psycholinguistic substrates of aspectual coercion, a linguistic computation whereby a temporally enriched interpretation of an event is produced in the absence of an overt syntactic marker. This typically takes the form of a mismatch between an aspectually marked verb and a temporal (aspectual) modifier. For example, “The street light flashed at dawn” is presumed to represent a single, transient event (i.e., the light flashing once), while “The street light flashed until dawn” represents a repetitive series of events. The latter interpretation of “flashed repeatedly” is obligated by the resolution of the mismatch between the point action verb (“flashed”) and a temporally durative adverb (“until”). This interpretation has been hypothesized to be compelled semantically, rather than syntactically, as the syntactic structure of both sentences is identical. Previous studies using cross modal lexical decision (Piñango, Zurif, & Jackendoff, 1999; Piñango, Winnick, Ullah, & Zurif, 2006), stop-making-sense tasks (Todorova, Straub, Badecker, & Frank, 2000a, 2000b), and self-paced reading (Proctor, Dickey, & Rips, 2004) have found behavioral evidence for increased processing load during the coercion process (but cf. Pickering, McElree, Frisson, Chen, & Traxler, 2006), but the precise time-course of its operation and its (psycho)linguistic nature have yet to be confirmed empirically. This paper presents an event-related potential (ERP) study designed to confirm the linguistic position that aspectual coercion is a (primarily) semantic operation and to attempt to resolve conflicting evidence regarding its time course. Results support the contention that aspectual coercion is, indeed, a purely semantic operation, indexed by an

increased N400 component in coerced conditions. Furthermore, the data suggest that effects of aspectual coercion operate differently when the coerced verbs are inherently temporally unbounded (e.g., “glide”) and when they are inherently point-action (e.g., “hop”). Together, these results indicate that aspectual coercion is a semantically-based process that generates a semantically enriched interpretation of event temporal structure, limited primarily to the creation of iterative interpretations but not merely durative, continuous ones (e.g., Todorova et al., 2000a).

3.1. INTRODUCTION

Psycholinguistic models of sentence level comprehension have typically focused on the building of syntax, or grammar (i.e., how the order of words tells the reader/listener who did what to whom) as an initial stage in the process leading toward ultimate interpretation (e.g., Clifton & Frazier, 1989). The process of aspectual coercion, however, imposes a temporally “enriched” interpretation – not instantiated in a purely structural, or syntactic, reading – upon an event. *Aspect* relates to the internal temporal structure of an event, typically indicating whether an event has been terminated (e.g., telic) or is an ongoing process (e.g., atelic). Verbal lexical aspect, a lexical property of verbs, may have a finite and bounded nature (e.g., hit, sneeze, kill, etc.), indicating that the event represented is a single event which was completed, or an indefinite and unbounded nature (e.g., love, want, search, etc.), in which no temporal boundary is presumed (as in Mourelatos, 1981)¹⁷.

On a sentential level, enriched interpretation due to aspectual coercion is not imposed by any individual lexical item in the sentence, nor is it overtly expressed in the verbs (in English¹⁸). An enriched interpretation is the result of the resolution of a mismatch between the aspect of an event (verb) phrase and its aspectually marked temporal modifier. When a durational modifier is matched with a verb with no inherently durational sense – as is the case with verbs representing single, punctual events – this typically leads to a sense of repetition. In this way, a distinction is made

¹⁷ For elaboration of these and other possible classifications of aspect, see Dowty (1978), Vendler, (1967).

¹⁸ ...Or other languages which differentiate morphosyntactically between the *perfective* and *imperfective* tenses, such as Spanish.

between *point-action* verbs, which, when under the influence of aspectual coercion, represent events occurring iteratively (e.g., a single event occurring over and over again), and *process* verbs, which represent events that occur in an ongoing, continuous manner, regardless of the temporal boundedness of the context.

Thus, Sentence (1a) describes an event that occurs once, while Sentence (1b) describes an event that occurs repeatedly over an extended period of time.

(1a) The street light flashed at dawn.

(1b) The street light flashed until dawn.

The difference in interpretation between (1a) and (1b) is due exclusively to the temporal influence of the adverbial preposition and its durative (aspectual) properties. The verb *flashed* is a point action verb, and has the sense of a single occurrence, e.g., of a single pulse of light (1a). In (1b), however, the duratively marked preposition *until* requires that the single event aspect of *flash* be revised, or *coerced*, into an iterative, repetitive interpretation, with the ultimate meaning that a series of flashing events takes place for some extended period of time, ending at dawn.

Sentences (2a) and (2b), on the other hand, are both events that take place continuously over time – a short (but continuous) period in (2a) and an indeterminately long (still continuous) period in (2b).

(2a) The street light glowed at dawn.

(2b) The street light glowed until dawn.

The verb *glowed* in (2a) and (2b) is a process verb, having an inherent, durative aspect which implies a continuous, constant state. Although process verbs may also represent temporally bounded, single-episode events (e.g., (2a) above), the

inherent logic of the verb conveys the sense that the light glows continuously, even if only for a brief time. Likewise, in (2b), the light clearly glows in an ongoing, uninterrupted manner (but not iteratively/repetitively).

Aspectual coercion is thus seen to be a primarily semantic operation in which the conceptual structure of an event or events is re-computed when the human language processor (i.e., the parser) resolves an aspectual (lexically-based) mismatch. In other words, aspectual lexical properties of the verb interact with lexical properties of the temporal modifier; such a mismatch must be resolved if the correct interpretation is to obtain. Specifically, when a punctual verb (e.g., *flash*) is operated upon by a durative adverbial modifier (e.g., *until*), an iterative interpretation must be computed. On the other hand, durative verbs (e.g., *glow*) in the same context may not require such computation, as they are inherently continuous.¹⁹ A number of studies have investigated the psycholinguistic nature and time course of this computation.

3.1.1. Processing costs

A small number of studies have investigated this particular version of aspectual coercion.²⁰ The first of these, Piñango, Zurif, and Jackendoff (1999), examined this

¹⁹ A typical example might be “*The street light glowed for a minute*”; even though it only glows for a short time, it is presumed to glow continuously throughout that time period. An obvious exception to this is a case in which a durative verb is operated upon by a modifier representing a very small temporal window, e.g., “*The street light glowed at exactly 7:59 and 12 seconds*”. While this might be considered a form of aspectual coercion, such a context is unlikely to be represented strongly in the corpus, and is therefore not under investigation in this paper.

²⁰ Another form of aspectual coercion, “complement coercion” (e.g., “The author began the book...” in which *began* refers to an entity rather than an event, leading to listeners coercing the complement into an event), has been studied extensively

issue using the cross modal lexical decision (CMLD) interference task. Participants heard sentences such as (3a&b) below and had to make lexical decisions to word strings which appeared on a computer screen at a critical time during the auditory stream. The

(3a) The boy screamed loudly until^ well into the evening...

(3b) The boy belched loudly until^ well into the evening...

In (3a), *screamed* is a temporally durative verb; it implies a continuous process. Because it also is matched with a durative adverbial modifier, it is therefore considered “transparent” with regards to its interpretation. That is, *screaming* is presumed to take a certain amount of time, and so when a temporal phrase like *until* is encountered, it does not require recomputation. On the other hand, (3b) contains the verb *belched*, which is punctual (i.e., point action). Upon encountering *until*, the single, point action interpretation of *belched* must be reinterpreted, and thus the semantic content of the verb (and its corresponding context) is “enriched” to having an iterative sense.

Piñango et al. (1999) found that lexical decision times made to items presented at 250 ms²¹ following the coercing temporal modifier (denoted by the caret above) were longer during enriched sentences (i.e., point action verbs coerced to be repetitive) than in sentences which already have a continuous interpretation (i.e., process verbs undergoing no coercion). Piñango et al. interpreted this increased processing load as

elsewhere, and will not be discussed here (e.g., McElree, Traxler, Pickering, Seely, & Jackendoff, 2001).

²¹ This specific number was chosen because it was between the 230 ms and 279 ms time points at which McElree and Griffith (1995), in an unrelated study, found reliable detections to syntactic and semantic violations, respectively.

the engagement of the process of aspectual shift; processing load specifically reflected the creation of an iterative sense. A follow-up questionnaire to a second group of participants demonstrated that this effect was not due to differences in plausibility between the two conditions (but see De Velle, 2003).²²

A recent follow-up study (Piñango, Winnick, Ullah, and Zurif, 2006) supported and somewhat extended their earlier results. Using the exact same stimuli as those in their earlier paper, Piñango and colleagues were able to replicate the results from their earlier study, showing that aspectual coercion exacts a processing cost when the probe was presented 250 ms following a temporal modifier that coerces an enriched interpretation. Despite identical stimuli, the RTs in their second study were considerably shorter than those in the original study (661 ms for transparent and 682 ms for enriched in the second study, compared to 743 ms and 782 ms, respectively, in the original study). To extend and elaborate their earlier results, Piñango et al. conducted a second experiment to determine whether aspectual coercion may be observed as soon as it is structurally licensed, i.e., *immediately* following the temporal modifier (as opposed to 250 ms after it). When the probes were placed immediately after the temporal modifier, the statistical analyses failed to find a significant difference between probe reaction times in the transparent and enriched conditions. In fact, reported t-tests comparing RTs to probes in transparent (710 ms, SD = 161)

²² The enriched verbs in the Piñango et al. (1999) study were more frequent, on average, than the transparent verbs. Because the authors suggested that more frequent verbs provide a lower computational cost than less frequent verbs, they concluded that that their bias in verb frequency (i.e., favoring the “enriched” verbs) could not explain the increase in reaction time (RT) for the lexical decisions made during enriched sentences.

versus enriched (722 ms, SD = 151) sentences revealed a $t(19) = 1.29$, $p = 0.104$. Hence, although the authors are correct in stating that these numbers fail to reach significance, their failure to acknowledge such a clear trend – especially given the large SDs in their data – is surprising. At the very least, such a trend invites the speculation that, although perhaps not exactly at the point at which it is syntactically licensed, but certainly shortly thereafter (probably before 250 ms), the influence of semantic enrichment probably takes effect.

Other authors have found evidence of increased processing demands at the point in sentences in which aspectual coercion is licensed. Using a stop-making-sense task, Todorova et al. (2000a) presented stimuli controlling cardinality of predicate objects (i.e., singular vs. plural) and durative vs. non-durative modifiers to coerce interpretations. Sample stimuli can be seen in Table 1. These authors examined reading times and sense rejections (the point at which the sentence “does not make sense”) in a self-paced, makes-sense judgment task, in which participants were instructed to read consecutive portions of sentences quickly, advancing the regions once they were read, and to indicate whether each portion made sense with respect to the previously presented material. They found increased reading times and increased sense rejections to portions of the sentences containing mismatching aspect, but only when durative modifiers were linked with singular predicates. These results are consistent with increased processing costs association with processing enriched interpretations. Based on their results, the authors claim that aspectual coercion is a costly operation, but the observed processing delay might reflect 1) a semantic reanalysis of an incorrect initial interpretation (formed based on the predicate before

the temporal modifier is reached), or 2) the costs of generating an iterative interpretation. Unfortunately, they are unable to differentiate these two based on their design.

Other authors, meanwhile, have found evidence that aspectual coercion exacts a processing cost but only late in the sentence, i.e., well after coercion is licensed. Proctor, Dickey, and Rips (2004) crossed verbs differing in telicity (e.g., “consumed” vs. “monitored”) and mass vs. count nouns (e.g., “ice water” vs. “ice cube”) to produce coerced interpretations in written sentences. In a segment-by-segment self-paced reading task, they found that participants postponed telicity judgments in coerced sentences until the last segment, rather than immediately once the aspectual information is processed. That is, participants did not show increased reading times in the third segment, which included the verb’s object (which provided aspectual information leading to coerced or non-coerced interpretations); they only showed increased reading times in the fifth, final segment, when the meaning of the entire sentence was available. This suggested to Proctor et al. (2004) that the processing of aspect may not be performed incrementally (minimal commitment; see below) but may be postponed until inferences about telicity must be made.

These previous studies have all found evidence that aspectual coercion requires increased processing, though they have varied with respect to the time at which that processing cost is exacted. One large set of experiments failed to find evidence of increased processing during aspectual coercion. Pickering, McElree, Frisson, Chen, and Traxler (2006) attempted to replicate both the Piñango et al. (1999) and Todorova et al. (2000) studies, re-running the actual stimuli from both those studies in both self-

paced reading and eye-tracking methodologies (i.e., Pickering et al. presented 4 experiments). An important feature of the Pickering et al. studies is that they did not require participants to perform a task *during* sentence processing; comprehension questions to the sentences were presented after trials²³, so as to permit uninterrupted processing.

In their series of 4 experiments failing to replicate Piñango et al. (1999) and Todorova et al. (2000), Pickering et al. (2006) claimed that:

Effects of aspectual coercion appear to engender measurable differences only when reading is paired with a concurrent task, such as a lexical-decision task (Piñango et al., 1999) or a stop-making-sense task (Todorova et al., 2000a, 2000b). We believe that these secondary tasks induce comprehenders to commit to properties of an interpretation that they would otherwise leave underspecified. (p. 150)

In effect, these authors claim that during “typical”, uninterrupted reading conditions, readers do not engage in a *full commitment* strategy (i.e., immediately committing to a minimal interpretation with each incremental portion of information), but rather postpone committing to a particular interpretation by using an *incomplete commitment* strategy, in which the representation of meaning as the sentence is incrementally processed remains underspecified until commitment to form an interpretation is necessary (e.g., for a comprehension question after the sentence). Although it is possible that task-based effects may result in differential processing, it is unclear how

²³ In the Pickering et al. Experiments 1 and 3, participants were required to answer questions after each sentence they read; in Experiment 2, they responded only after “some” of the filler sentences; in Experiment 4, no details regarding comprehension questions are provided. Even though the authors suggest that the comprehension questions (in Experiments 1-3) would encourage participants to process sentences deeply, according to Pickering et al., the questions were “relatively straightforward and did not particularly focus on the aspectual coercion itself” (p.138).

comprehenders should know whether any particular portion of a sentence will require postponing commitment. Thus, if participants may not even perceive that a different interpretation needs to be made until forced to by a secondary task. Combined with the non-intuitive nature of this hypothesis, and compounded by the null-result basis of their studies, this explanation is particularly unsatisfying.

Indeed, Pickering et al. (2006) suggest that aspectual underspecification may be under a certain amount of strategic control, and that individual readers may be sensitive to the amount of processing resources required for any given passage and adjust their parsing strategies accordingly. Such a strategic postponement of commitment would ostensibly require a certain amount of processing resources in order to maintain all the possible interpretations “in mind” until disambiguating information is presented. Though this introduces an intriguing set of theoretical questions dealing with individual differences in language processing (e.g., working memory constraints), a deep consideration into the substantial literature on processing resources is beyond the scope of this paper. It is furthermore not unreasonable to suspect that strategic consideration of processing demands should, in and of itself, result in increased processing costs beginning at the point of ambiguity and “releasing” at or shortly after the point of disambiguation. Pickering et al.’s failure to find such a processing difference suggests that their explanation can not fully account for the findings of the other studies.

One curious feature of the Pickering et al. (2006) paper is the authors’ choices in data management. For example, their replication of Piñango et al. (1999) in Experiment 1 provided reading time data for each word (and some short phrases) in

the sentences. The authors state that responses less than 200 ms were recoded as 200 ms, and responses greater than 2000 ms were recoded as 2000 ms. (In Experiment 3, these cutoffs were set to 100 ms and 4000 ms, respectively, to accommodate the increased length of visual units, i.e., (mostly) two- to three-word chunks instead of single words.) It is not clear why the investigators chose these numbers, as they do not comport with common practice of excluding data > 2 SD above/below the mean, nor do the authors indicate – given the potentially skewed RTs imposed by these cutoffs – whether the response distributions were sufficiently normally distributed to justify subjecting them to ANOVA.

To summarize, Piñango et al. (1999, 2006) provide evidence that aspectual coercion is computed almost immediately after the aspectual mismatch is licensed (250 ms after a temporal modifier), and Todorova et al. (2000) found increased processing at and immediately following a region containing a temporal modifier licensing coercion. Proctor et al. (2004) found that self-paced readers only show increased processing times at the final sentence segment, after all the aspectual information has been presented, whereas Pickering et al. (2006) failed to detect processing costs associated with aspectual coercion in any of their 4 experiments (2 reading time, 2 self-paced reading).

Despite possible methodological shortcomings, the Pickering et al. results present an intriguing possibility – that depending on task demands, comprehenders may not pursue an immediate (full commitment) interpretation. Specifically, comprehenders may not compute aspectual mismatches of the type presented by Piñango et al. (1999, 2006) and Todorova (2000a, 2000b) when they are engaging in

“typical” on-line processing. It is widely accepted that methodology choices in psycholinguistic paradigms can impart a certain amount of bias due to the type of processing participants must engage in during the task (Swinney, 2000; Love & Swinney, 1996; Nicol & Swinney, 2003; Nicol, Swinney, Love & Hald, 1997; Chwilla, Brown, & Hagoort., 1995, among others). This issue raises the clear need for an augmentative methodology to address this question; a methodology to visualize processing costs associated with aspectual coercion without the need for a secondary task. ERPs appear to be appropriate for addressing such linguistic processes.

3.1.2. Event-related potentials

Event-related potentials measure minute fluctuations in the electrical activity produced by cortical neurons in the brain, time-linked to specific cognitive processes. Numerous studies have supported the sensitivity of ERPs to specific linguistic information, particularly with regard to semantic integration versus syntactic processes (Kutas, Federmeier, Coulson, King, & Münte, 2000; Kutas, Van Petten, & Kluender, 2006). In general, difficulty integrating semantic information into the ongoing context results in an increase in a negative deflection peaking around 400 ms after the point of integration (i.e., the N400 component; Kutas & Hillyard, 1980, 1983; Kutas & Van Petten & Kutas, 1994; Friederici, Steinhauer, & Frisch, 1999; Hagoort & Brown, 2000). Based on its sensitivity to semantic, conceptual processes, the N400 seems the likeliest candidate for indexing increased processing loads implicated in aspectual coercion.

Another component, the P600/LPC, also holds potential for indexing aspectual coercion. Although it is generally considered an index of syntactic violation or reinterpretation (Osterhout & Holcomb, 1995; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992; Hagoort, Brown, & Osterhout, 1999), the P600 has been shown to index structural complexity or syntactic integration (e.g., Kaan, 2000), and has been observed in situations requiring sentence/clause reanalysis following syntactic and/or semantic mismatches (Münte, Heinze, Matzke, Wieringa, & Johannes, 1999; Friederici & Mecklinger, 1996; and many others). If the linguistic operation of aspectual coercion involves resolving a mismatch in lexical aspect, the P600 may index increased processing load related to structural reanalysis. If aspectual coercion is, indeed, a purely semantic operation, however, only the N400 would be expected to index the increased processing costs of coercion.

The current study was undertaken to determine whether ERPs might provide insight into the linguistic nature of aspectual coercion, and to attempt to further delineate the time-course of its operation. Several hypotheses may be generated, based on the results of behavioral studies, and based on the broad assumption that the linguistic system should process semantic information quickly, on-line. First, aspectual coercion is hypothesized to operate semantically. When point action verbs are coerced into an iterative, repetitive interpretation, this conceptual switch should be indexed by a greater N400 for the coerced (or “semantically enriched”) condition than for the un-coerced (or “transparent”) condition. An N400 effect is expected relatively immediately following the processing of the coercive temporal modifier. If increased processing during aspectual coercion is indeed the result of generating an iterative

interpretation, then an N400 effect should be seen for point action verbs being coerced, and no such N400 effect should be observed when process verbs are processed in the same coercing contexts. That is, process verbs linked to both punctual and durative modifiers should be interpreted as being continuous (and not iterative), and so comparison between process verbs should show no differences in the ERP (i.e., no N400 effect for “semantically enriched” compared to “transparent” interpretations in process verbs) when matched to *at* vs. *until* modifiers. Hypothesis 1 can be summarized as follows: Point action verbs being aspectually coerced should show an N400 effect, with enriched interpretations giving a more pronounced N400 than transparent interpretations; process verbs in the same contexts should produce no such N400 effect. Identical N400 effects for both verb types would indicate that the operation of aspectual coercion generates

Secondly, direct comparison between iteratively enriched point action verbs and duratively transparent process verbs will demonstrate whether the psycholinguistic nature of both processes is identical. Enrichment in point action verbs is purported to generate an iterative interpretation, while both enriched and transparent process verbs both generate continuous (but not iterative) interpretations. Therefore, Hypothesis 2 predicts an interaction between coercion status and verb type, such that ERPs elicited by enriched vs. transparent point action verbs will differ, but ERPs elicited by enriched vs. transparent process verbs will not.

Finally, the Piñango et al. studies (1999; 2006) conflated two types of temporal modifier phrases; this was presumably done under the assumption that the coercion operation proceeds equally under both modifier contexts. In addition to the *at* vs. *until*

modifier type mentioned earlier, Piñango et al. used a prepositional phrase, having the form of “*for [temporal modifier]*”:

(4a) The man jumped for an hour.

(4b) The man swam for an hour.

The interpretation of (4a) is that of an iterative event; due to the conventional representational properties of *jump*, a man cannot engage in a single jump lasting an hour. Thus, he must jump repetitively over the course of the hour. This interpretation is aspectually coerced based on the properties of the duratively marked prepositional phrase. This structure was matched in the Piñango et al. stimuli with durative (process) verbs, as in (4b). Their use of this additional type of modifier phrase presumes identical processing between the two structures; a cursory glance at the structure of the modifier, however, indicates that this may not be the case. Time-locking to the object of the preposition (which is presumably the point which licenses the aspectual interpretation) may miss the point at which listeners actually process or perceive that a coercion operation is required. That is, listeners may be sensitive to the fact that hearing a phrase beginning with “*for the...*” tends to end in a durative temporal modifier, particularly when confronted with a limited variety of experimental stimuli such as those used by Piñango et al. In this way, listeners in their experiments might have engaged in predictive strategies, committing earlier in those prepositional stimuli than in stimuli in which the durative modifier is the adverbial *until*. Thus, another version of the prepositional modifier was included in the present study, one which leads to an explicitly non-durative interpretation.

(4c) The man jumped for an apple.

In (4c), a single jump is made to reach the apple. Whether the apple is actually reached is not expressed, but this interpretation permits the single event, point action nature of *jumped* to be matched with an object or theme rather than a temporal modifier; no coercion is presumed to occur in such a context. Furthermore, no mismatch occurs even when used in the context of an inherently durative, process verb, like *sang* in (4d&e) below. As mentioned above, verbs with imperfective aspect can be used to generate a bounded, continuous interpretation, as in (4c), or a temporally unbounded (but non-durative) interpretation, as in (4e).

(4d) The man sang for an hour.

(4e) The man sang for an award.

Because Piñango et al. (1999, 2006) appear to have conflated²⁴ the two types of durational modifiers (*when/until* vs. *for an hour*), this study included a paired set of sentence stimuli having modifiers of the form “*for...*” with both durational and non-durational modifiers, in order to 1) determine whether these types of modifiers result in different processing, and 2) reduce the likelihood of strategic processing leading to varied time-courses of coerced interpretation commitment.

In summary, most empirical results to date suggest that aspectual coercion involves a “processing cost” of some kind. Increased lexical decision times (Piñango et al., 1999; 2006), and increased reading times (Proctor, Dickey & Rips, 2004; Todorova et al., 2000a) have been used as evidence that some form of mismatch is being computed between aspectual properties of a verb and a (temporal) predicate

²⁴ Of their 25 sentence pairs, Piñango et al. (1999, 2006) used the following temporal modifiers (number of times in parentheses: “until” (13), “for an hour” (5), “for a long time” (3), “for hours” (2), “for almost an hour” (1), and “for about two hours” (1).

operator (cf. Pickering et al., 2006). The precise time course and linguistic nature of such processing costs, however, are unclear. These studies suggest commitment to a particular sense of the verb has been made – immediately after it is licensed in Todorova et al. (2000), shortly thereafter (Piñango et al., 1999, 2006), or at the end of the sentence (Proctor et al., 2004). Thus, ERPs were used in the present study to elaborate on the time course of aspectual coercion.

3.2. METHOD

3.2.1. Participants

Twenty-five (25) participants took part in this study. Participants were students at UCSD, who received course credit for their participation. Data from 4 participants were discarded due to an excessive number of movement or eye-blink artifact-influenced trials (> 30%); 3 participants' data were discarded due to experimenter error (recording problems), and 3 participants' data were discarded due to technical issues (corrupted data files). Consequently, data from nineteen (19) participants (5 male, 14 female) are included in the final analyses. Mean age of participants was 20.6 years (range 18-25, SD = 1.9). All were right-handed as assessed by a modified Edinburgh handedness inventory (Oldfield, 1971), native speakers of English with no exposure to another language before age 6, and no visual, auditory, or neurological impairments.

3.2.2. Stimuli

Experimental items consisted of 80 pairs of English sentences: 40 pairs contained point action verbs (e.g., *flash*) and 40 pairs contained process verbs (e.g., *run*), organized into minimal pairs. One set of pairs examined point action verbs in transparent vs. enriched conditions, and the other set examined process verbs, transparent vs. enriched. Of the 40 pairs for each verb type, one half (20 pairs) contained temporal modifiers of the “For...” type and half of the modifiers were of the *at/until* type (20 pairs each for point action and process verbs).

To summarize the different conditions, a list of abbreviated example sentences appears below (5a-h); the final set of 80 experimental items was comprised of 10 items from each condition. The gloss below each item describes its verb type (point action vs. process), coercion status (enriched vs. transparent), and modifier type (“For...” vs. *at/until*), respectively. The caret (^) represents the point to which ERPs were time-locked (i.e., the point at which coercion is licensed):

- (5a) The boxer punched his opponent after^ he became tired...
(Point action, transparent, *at/until*)
- (5b) The boxer punched his opponent until^ he became tired...
(Point action, enriched, *at/until*)
- (5c) The criminal watched the witness before^ she left...
(Process, transparent, *at/until*)
- (5d) The criminal watched the witness until^ she left...
(Process, enriched, *at/until*)
- (5e) The girl dove into the pool for a penny^ with her mom watching...
(Point action, transparent, “For...”)

(5f) The girl dove into the pool for an hour[^] with her mom watching...

(Point action, enriched, “For...”)

(5g) The woman serenaded her lover for attention[^] before giving up...

(Process, transparent, “For...”)

(5h) The woman serenaded her lover for an hour[^] before giving up...

(Process, enriched, “For...”)

Two sets of pre-tests ensured that the sentences 1) expressed the intended temporal interpretation, and 2) were equally matched for plausibility. Pre-test participants were undergraduate students at the University of California, San Diego (UCSD) who received course credit for their participation; pre-tests were administered in conjunction with offlines for other unrelated experiments.

Pre-test 1:

A total of 28 participants (13 male, 15 female; mean age = 19.7, SD = 1.3) read pseudo-randomized lists containing 42 candidate sentence pairs. Each sentence pair consisted of one “repetitive” interpretation and one “single time” interpretation, all of which used the “for...” modifier; coercion in “enriched” versions was due to a temporal component after “for”, while punctual interpretation in “transparent” versions was conveyed by an agentive or thematic object. Participants were asked to rate each sentence on how plausible it felt and whether it indicated a repetitive/continuous action, each on a scale from 1 to 9, with “1” representing very implausible/non-repetitive and “9” representing very plausible/repetitive. Mean ratings were compared across repetitive and punctual interpretations. A total of 22 matched sentence pairs were rated according to the intended meaning (determined by

a difference in mean rating of 4 units, with repetitive sentences rated greater than 4 units more repetitive than single-action sentences). The rest of the sentences were re-structured²⁵ and administered in pre-test 2.

Pre-test 2:

Nineteen (19) participants (7 male, 12 female; mean age = 20.3, SD = 1.7) read pseudo-randomized lists containing 63 candidate sentence pairs, 42 of which used the *until/when* adverbial modifier and 21 of which used the “for...” durational modifier. Again, only matched pairs whose average repetitiveness ratings differed by more than 4 units were kept.

The final sentence list of 80 experimental items was comprised of the “best” items (i.e., greatest difference between repetitiveness ratings while having equivalent plausibility ratings) taken from across both pre-tests: 20 sentence pairs using *at/until* and 20 using “for...” modifiers. Eighty filler stimuli were made up of sentence pairs having a variety of structures to reduce pattern learning and strategic processing in participants.

Items were digitized on a computer using CoolEdit Pro v1.2 by a female, native English speaker with a standard American English dialect. Each trial began with “Ready”, followed by a 1000 ms pause, followed by the sentence. Each sentence was followed by a 1500 ms silence, at which time an auditory tone signaled time for response.

²⁵ In cases where the desired interpretations could not be achieved, the sentences were re-designed with the *until/before* modifiers and included in pre-test 2.

3.2.3. Procedure

Participants were seated comfortably in a sound-proof booth approximately 1.65 meters from a pair of stereo speakers. Participants were instructed to fixate their eyes on a point on a monitor (also 1.65 m) as they listened to stimuli. They made a binary, button-press judgment after each item depending on their judgment of whether the sentences represented actions which occurred once or repeatedly/continuously. Because of the importance of making certain that participants understood the task, several examples were given to ensure that the participants understood how to respond correctly.²⁶ Participants were asked to remain still without blinking while listening to the entire sentence and to press the button only after they heard the beep.

Because of the number of trials, participants came in for 2 visits, no less than a week apart. Four lists of items were pseudo-randomized (such that no more than two items from the same condition were heard consecutively), having the following order (where AB and CD each represents a generic block of 20 items, and BA and DC represent those blocks presented in reverse order): ABCD, BADC, CDAB, DCAB. At each visit, participants heard 80 sentences, half of which were fillers, and half of which were experimental sentences containing point action and process verbs (10

²⁶ Example items given during instruction were:

Repetitive: “The horse galloped through the meadow.”

Continuous: “The boy reads the book.”

Single time: “The boy dropped the pencil.”

Participants were then asked to judge the following sentences, and were given feedback to ensure they understood the task completely:

“The man caught the fish.” (single time)

“The women fell into the ocean.” (single time)

“The ball bounced down the hill.” (repetitive)

“The man watched television.” (continuous)

enriched and 10 transparent, each, per visit). In their second visit, participants heard the matched sentences in “mirror image” order, such that an enriched version of a sentence in visit 1 was heard as its transparent version in visit 2, and vice versa. Response hand for “single time” versus “repetitive/continuous” was counterbalanced across participants. Total time for each visit (including set-up, testing, and disassembly) was ~60 minutes.

3.2.4. ERP procedure

ERPs were collected using an electrode cap (QuikCap, CompuMedics, Inc.) embedded with 32 sintered Ag/AgCl electrodes applied to the scalp, with all electrodes referenced to the left mastoid. Electrode placement was based on the International 10-20 System (Jasper, 1958), with additional sides: Fz, FCz, Cz, CPz, Pz, Oz, FP1/2, F3/4, F7/8, C3/4, T7/8, P3/4, P7/8, O1/2, FC3/4, FT7/8, CP3/4, and TP7/8. In addition, bipolar electrodes were applied to the outer canthus of the both eyes and above and below the left eye to monitor vertical (VEOG) and horizontal (HEOG) eye movements, respectively.

The EEG signal was amplified 100 X with a SynAmps differential amplifier (3dB cutoff, bandpass filtered on-line between 0.15 and 30 Hz), digitized on-line at 500 Hz. Average ERPs were formed off-line from trials free of EOG and movement artifact ($>\pm 100\mu\text{V}$).

3.2.5. Data analysis

A total of 5 participants exhibited excessive amounts of artifact (~50% of epochs were unusable due to movement and/or eye blinks) and their data were excluded from later analysis. One of the remaining participants showed a low accuracy rate (<70% correct) in the behavioral responses and was therefore excluded; all other participants performed >80% accuracy.

The ERP data to the conditions of interest were quantified by measuring the mean amplitude (in μV) within three latency windows determined to be, *a priori*, most appropriate for investigating the predicted effects: 100-300 ms to measure early effects, 300-500 ms to measure N400-related effects, and 500-1000 ms to investigate possible P600 effects. Grand average analyses were performed separately for each time window, compared to a -100 to 0 ms baseline. Repeated measures omnibus ANOVA (rmANOVA) was used for initial 2 X 2 X 2 analysis, with the following within-subject factors: Verb Type (Point Action vs. Process), Coercion Status (Transparent vs. Enriched), and Modifier type (*at/until* vs. “for an hour/apple”). The within-subject factors were crossed with Grouping variables: electrode (28 levels); laterality (Left vs. Midline vs. Right), anteriority (Anterior: F3/4, Fz, FC3/4, FCz; Centroparietal: Cz, C3/4, CPz, CP3/4, Pz, P3/4). Main effects or interactions in the omnibus rmANOVA led to factor-specific follow-up ANOVAs. The Huynh-Feldt correction (HF) was imposed to reduce likelihood of Type I errors when degrees of freedom in the numerator was greater than 1. All reported *p* values are HF corrected.

All analyses were conducted on unfiltered data. In the figures, waves have been subjected to a 0.3 – 9 Hz bandpass filter for visualization purposes.

3.3. RESULTS

Omnibus rmANOVA with Verb type (2 levels) X Coercion (2 levels) X Modifier type (2 levels), crossed with Electrode (28 levels), are summarized in Table 2. Results are presented for each of the 3 analysis windows.

3.3.1. Verb Type

No main effect of verb type (all p s $> .1$) was seen in any window (Figure 1). Two-factor ANOVAs showed no effects of Laterality or Anteriority on verb type in any time window (all p s $> .1$). Surprisingly, the interaction between verb type and coercion was not significant in any of the windows; Figure 2 illustrates this comparison. Although an interaction was hypothesized, failure to find a significant difference may indicate that both types of verbs are equally affected by the operation of aspectual coercion. This possibility is elaborated in the Discussion, below.

An interaction between verb type and modifier type was also found (described in the Modifier analyses below).

To determine whether the effects of coercion on verb type were indeed non-significant, and to further test planned comparisons, separate analyses were run for each verb type.

Point Action Verbs

When point action verbs were taken as a whole, an effect of modifier was found in all 3 windows (100-300 ms: $[F(1, 18) = 17.459, p = 0.0006]$; 300-500 ms $[F(1, 18) = 20.067, p = 0.0003]$; 500-1000 ms $[F(1, 18) = 32.126, p < 0.0001]$, with *at/until* type modifiers were significantly more negative than “For...” type modifiers.

Also, when point action verbs were taken as a whole, an interaction between coercion and modifier approached significance in the 100-300 ms window [$F(1, 18) = 3.386, p = 0.0823$], becoming significant in the 300-500 ms [$F(1, 18) = 5.947, p = 0.0253$] and 500-1000 ms [$F(1, 18) = 10.448, p = 0.0046$] windows. This took the form of enriched sentences have more negative voltages than transparent contexts when the modifier was of the “For...” type but the opposite difference (transparent contexts more negative than enriched) when the modifier was of the *at/until* type. If enrichment is taken to be indexed by a negative deflection in the ERP (N400-type process), then this result suggests that sentences with the “For...” modifier were possibly more successful in coercing an aspectual interpretation than sentences with the *at/until* modifiers.

Surprisingly, rmANOVA revealed no main effect of coercion in point action verbs in any time window (all p s > .1).

Process Verbs:

Even more surprising than the lack of a main effect of coercion in point action verbs (given hypotheses concerning the continuous nature of process verbs) was the finding that, when process verbs were taken as a whole, coercion exhibited a main effect on process verbs in all 3 time windows (100-300 ms: [$F(1, 18) = 5.771, p = 0.0273$]; 300-500 ms: [$F(1, 18) = 9.815, p = 0.0057$]; 500-1000 ms: [$F(1, 18) = 7.694, p = 0.0125$]), such that enriched process verbs were reliably more negative, overall, than transparent process verbs, across the entire epoch. Based on the amorphous morphology of this negative drift, however, it is unclear what type of process it represents; see Discussion.

No effect of modifier type was found in any time windows within process verbs (all p s > 0.05), although an interaction between coercion and modifier type approached significance in the 500-1000 ms window [$F(1, 18) = 3.759, p = 0.0684$], with enriched sentences showing more negative amplitudes than transparent sentences, but only for the “For...” modifier type; enriched and transparent sentences in the *at/until* sentences did not differ from each other in amplitude.

Coercion Context:

Figure 3 displays ERPs reflecting the overall effects of coercion context. A main effect of coercion became significant only in the late, 500-1000 ms window [$F(1, 18) = 6.742, p = 0.0182$]; coerced (enriched) interpretations produced deflections that were overall more negative than non-coerced (transparent) interpretations.

3.3.2. Modifier Type

A main effect of modifier type was found in all 3 time windows (100-300 ms [$F(1, 18) = 11.423, p = 0.0033$]; 300-500 ms [$F(1, 18) = 23.108, p = 0.0001$]; and 500-1000 ms [$F(1, 18) = 15.991, p = 0.0008$]), as shown in Figure 4. “For...” type modifiers resulted in ERPs that were more positive, overall, than *at/until* type modifiers across the entire duration of the epoch examined. The effect is consistent in scalp distribution (e.g., centroparietal) with a P600, but its very early appearance and early “peak” at ~300 ms indicates that it may not reflect processes associated with the traditional P600. This overall effect does suggest, however, that the two modifier types are being processed differently in some way, differentiating themselves at a very early time. This was confirmed (see Figure 5) by a significant modifier X verb type

interaction in the 100-300 ms [$F(1, 18) = 5.325, p = 0.0331$] and 500-1000 ms [$F(1, 18) = 5.994, p = 0.0248$] windows.

Because of the pervasive differences between modifier type, separate rmANOVA analyses were run within each modifier type. These results are also summarized in Table 4.

“For...” Modifiers

When “For...” type modifiers were taken as a whole, no main effects of verb type were seen (all $ps > .05$). There was a main effect of coercion in all 3 time windows [100-300 ms: $F(1, 18) = 3.912, p = .0635$; 300-500 ms: $F(1, 18) = 8.832, p = 0.0082$; 500-1000 ms: $F(1, 18) = 19.798, p = 0.0003$], with enriched (coerced) interpretations having a more negative deflection than transparent interpretations across the entire epoch, trending toward significance in the 100-300 ms window and becoming significant in the 300-500 ms window. Further ANOVAs of coercion within “For...” modifiers affecting point action verbs found effects of coercion only in the 500-1000 ms window [$F(1, 18) = 13.141, p = 0.0019$]. Main effects of coercion were also found with “For...” modifiers in process verb sentences in the 300-500 ms [$F(1, 18) = 7.715, p = 0.0124$] and 500-1000 ms [$F(1, 18) = 8.886, p = 0.0080$] windows (see Figure 6). These results indicate that the “For...” type of modifier was somewhat successful at producing a difference in processing between a durative (“For an hour”) and a non-durative (“For an apple”) interpretation, particularly in process verbs, reflected by a negative deflection during coerced interpretations.

At/Until Modifiers

When *at/until* type modifiers were taken as a whole, a main effect of verb type was found in the 500-1000 ms window [$F(1, 18) = 7.028, p = 0.0163$], with a trend toward significance in the 100-300 ms window [$F(1, 18) = 3.358, p = 0.0835$]. This difference revealed that point action verbs were more negative than process verbs. A borderline verb type X coercion interaction was found in the 300-500 ms window [$F(1, 18) = 3.298, p = 0.0861$], with transparent point action verbs being considerably more negative than transparent process verbs, while enriched point action verbs were more positive than enriched process verbs.

3.4. DISCUSSION

This study had several aims, the primary of which was to demonstrate that the “semantic sensitivity” of event-related potentials could be used to corroborate the semantic nature of aspectual coercion, the process whereby a (presumably semantic) mismatch between an aspectually marked predicate and a temporal modifier is resolved by the generation of a new, semantically enriched temporal interpretation. In the type of aspectual coercion examined here, the question was whether ERPs measured at the point at which coercion is licensed could index the semantic operation of aspectual coercion. To study this operation, sentences with point action verbs were matched to temporally durative modifiers (e.g., *until* as opposed to *at*), which should obligatorily coerce them into an iterative, or repetitive interpretation that is not overtly expressed in the morphosyntax. This generation of an iterative interpretation was hypothesized to be the source of increased processing costs sometimes found at the point of coercion in psycholinguistic studies (e.g., Piñango et al., 1999, 2006;

Todorova et al., 2000). As a control for this comparison, process verbs were also examined, with the hypothesis that they already inherently possessed the sense of a continuous, durative process, and as such would not engage the generation of an iterative interpretation. Evidence of coerced processing was hypothesized to come in the form of an increased N400-like effect for enriched compared to transparent contexts in point action verbs and not in process verbs; if the coercion process involves the (semantic) generation of an iterative interpretation, it was predicted to result in an N400 effect for enrichment in sentences comprised of point action verbs but not in those comprised of process verbs. Furthermore, assuming that the lexical property of aspect inherent in a duratively marked adverbial modifier (e.g., *until*) should not be any different than aspect carried by a temporally durative prepositional phrase (e.g., “for an hour), the type of modifier was hypothesized to be irrelevant to the success or failure of the aspectual coercion operation.

Results found that point action verbs and process verbs were not indexed differently, overall, in the context of auditory sentence comprehension, even when participants were given the task of determining the telic or atelic nature of the events represented in the sentences. Though perhaps not theoretically controversial, this result (albeit null) makes it unlikely that rest of the effects seen in the data were simply the result of differences in the way the two types of verbs are processed in general. It could have been the case that the two verb types were processed differently overall (based on mental representations of lexical properties associated with the verb classes), but this did not appear to be the case with these stimuli.

Overall, the ERP exhibited a long, continuous negative shift in contexts in which aspectual coercion was licensed (enriched), relative to contexts in which no coercion was intended (transparent). The centroparietal distribution of this deflection, coupled with its time-course, suggested that it may reflect an N400-like semantic integration process (Kutas & Van Petten, 1994), rather than a dependency-type or morphosyntactic relationship (which would have a leftward, anterior distribution consistent with LAN; King & Kutas, 1991; Coulson et al., 1995). This result supports the hypothesis that ERPs are sensitive enough to index (at least some) semantic aspects of the coercion operation.

Contrary to predictions that generating an iterative interpretation involves different processing than generating a continuous interpretation, a verb type x coercion interaction did not prove significant. This might be taken as evidence that a similar type of processing occurred during the two processes (or, more accurately, it may be taken as lack of evidence for distinct processing). A visual inspection of the waveforms, however, suggests what appears to be a sharper deflection in the point action condition (Figure 2). Despite statistically non-different mean amplitudes across the 3 time windows, the relative morphology between the ERPs from the two verb types, and their different scalp distributions, indicate that the coercion computation between the two verb types may indeed have been of a qualitatively different nature; specifically, an early deflection in coerced point action verbs appeared to have a sharper, more defined peak (confined to typical N400 scalp areas), while the lasting negative deflection in “coerced” process verbs was consistent, relatively undifferentiated, and broadly distributed.

Although not statistically compelling, a trend toward significance for this interaction in the 300-500 ms window hints that some type of operation, computed differently between the two verb types, may become evident during this time window but fail to be differentiable in the early and late windows. Indeed, it is possible that verb type and coercion contexts failed to show evidence of interaction in the earliest (100-300 ms) window because the main effects of coercion had not been processed fully at that time (cf. Piñango et al., 2006). Furthermore, an apparently similar profile (i.e., similar differences between coercion contexts across the two verb types) in the late window could reasonably be the result of the same general semantic operation being computed in both cases. In other words, the index of one semantic operation (e.g., generating an iterative interpretation) might be superimposed upon the index of another semantic operation (e.g., processing a mismatch in aspect), showing up as a borderline difference in only one of the time windows (specifically the point action verb type), leading to ERPs that are otherwise identical except for a peak “riding” upon a baseline “enrichment” shift. No differences were found in the overall comparison between point action and process verbs (see above), suggesting that this effect was not merely due to specific stimulus properties (e.g., that point action verbs had some physical quality – such as length or frequency – that interacted in a way that process verbs do not, or vice versa). This suggests that the (borderline) interaction across coercion conditions in this middle latency window may have indeed been due to the manifestation of neurally distinct computations between the verb types. This possibility is consistent with one potential set of hypotheses; i.e., that coercion affects both process and point action verbs, but that it does so using two distinct semantic

operations depending on verb type. For example, general effects of semantic enrichment might manifest across both verb types as a sustained negativity (e.g., if it indexes simply detection of an aspectual mismatch, or if it indexes a general interpretation of continuousness rather than punctuality), but coercion involving the generation of an iterative interpretation – elicited only by point action verbs – might result in a more sharply defined negative peak. The nature of this operation carries a sense of semantic reinterpretation; it may therefore manifest in a deflection having characteristics similar to N400 effects elaborated by Coulson and others (e.g., Coulson, 2001) resulting from “frame shifts” during, for example, comprehension of jokes and determining meaning from metaphorical utterances. Although this is purely conjecture (given that the present data do not specifically address this issue), this possibility would support the hypothesis that behaviorally demonstrated effects of aspectual coercion may be the result of the generation of an iterative interpretation for verbs with punctual aspect, above and beyond the processing of an aspectual mismatch (which is actually shared by both verb types). A finer grained analysis of the ERP in this window (e.g., narrower time periods using 50 ms windows) may permit a more complete elaboration of the effect.

When verb types were treated separately, only process verbs showed main effects of coercion, with coerced (enriched) interpretations overall producing greater negativity than non-coerced (transparent) interpretations. Process verbs were not predicted to permit iterative interpretations, so this generalized negative shift appears to be the result of some linguistic process resulting from the difference between telic (having a termination point, or perfective) and atelic (imperfective) interpretations.

Only point action verbs showed an interaction between coercion and modifier type. This latter result reflected relatively negative amplitudes to coercing (enriched) contexts than to non-coercing (transparent) contexts when the modifier was of the “For...” type (more negative when an enriched interpretation is generated); the opposite relationship was found when the modifier is of the *at/until* type (see Figure 6). If aspectual coercion is assumed to be indexed by greater negativity in enriched vs. transparent contexts, then it appears to be the case that “For...” type modifiers were able to generate coerced interpretations, while *at/until* type modifiers fail to produce coerced interpretations in point action verbs.

The overall effect of modifier type is particularly important and merits discussion. Results of this study indicate that aspect inherent to prepositional modifiers (*at/until*) appears to be processed significantly differently than aspect conveyed by adverbial modifiers (“For...”). A few possible explanations for this outcome are worthy of consideration. First, the grammatical category of the word which licenses coercion (or obligates non-coercion) was different across the two conditions, with the key word in *at/until* modifiers being a preposition and the key word in the “For an hour/apple” modifiers being a noun. This is a non-trivial distinction, particularly regarding ERP correlates of processing. Although there is some dispute about the exact time-course and distribution of the so-called N280 (typically a left, anterior negativity peaking around 280 ms elicited by closed class, “function” words such as prepositions, pronouns, and determiners), there is general consensus that closed class words produce different ERP signatures than open class words (Kutas & Hillyard, 1983; Van Petten & Kutas, 1991; Neville, Mills, & Lawson,

1992; Brown, Hagoort, & ter Keurs, 1999; Münte, Wierenga, Weyerts, Szentuki, Matzke, & Johannes, 2001; among others). It is therefore not surprising that the main effect of modifier type presented as an early negative peak (100-300 ms window, consistent with an N280 except for its centroparietal distribution) followed by a sustained negativity for *at/until* (closed class) modifiers, compared to “For...” (open class) modifiers. This effect appears to be driven by an early positive deflection in the “For...” modifiers (see Figure 6). Because this type of modifier appears to be the more appropriate context for finding coercion effects (see discussion on possible confounds in the *at/until* stimuli), an examination of the effects in only this modifier type can yield informative results.

Within “For...” modifiers, an effect of coercion was found across all three time windows, indicating that coerced (enriched) interpretations produced a generalized negative shift compared to transparent interpretations. (Within *at/until* modifiers, no effects of coercion were seen.) This may have to do with a confound imposed by the specific prepositions used. In the majority of the comparisons, *until* was used to coerce a semantically enriched interpretation compared to a punctual preposition such as *at* (see Appendix A), at least for point action verbs. However, the “non-coercive” preposition was at times inadvertently potentially identical to *until* in its durative aspect. For example, “The old man slept in the cabin until his friend came home...” may have the exact same aspectual sense as “The old man slept in the cabin before his friend came home.” Both can be taken to mean that the old man was sleeping and that the sleeping came to an end when the friend came home. Because the number of such

pairs was not controlled across the two verb types²⁷, it is impossible to be certain whether this had an effect on the overall “coercion-ability” of *at/until* modifiers.

Nevertheless, taken as a whole, these results suggest that the operation of aspectual coercion does, in fact, take place via a semantic computation, as indexed by a negative shift in the ERP trace for enriched vs. transparent contexts. This negative shift begins very early (as late as 100-300 ms following word onset), indicating that the semantic system is sensitive to the coercion operation almost immediately following the point at which it is licensed. This supports models which claim that the parser operates under a full commitment strategy, such that the lexicosemantic system is sensitive to mismatches between aspect in sentential event structure and a corresponding temporal modifier as soon as the coercing item is encountered during on-line sentence processing (Piñango et al., 1999, 2006; Todorova et al., 2000). These results provide strong evidence against models which posit that the parser engages in noncommittal or delay strategies, only computing coercion operations when necessitated by a secondary task (e.g. Pickering et al., 2006; Proctor, Walsh, & Rips, 2004).

These results also suggest that the time-course of coercion computation may be considerably earlier than previously described in studies using cross modal lexical decision (Piñango et al., 2006) and phrase-by-phrase self-paced reading (Todorova et al., 2000) tasks. Indeed, the present results using the ERP methodology provide a considerably more direct window into the timing of this process, as no secondary task

²⁷ Point action verbs: “until” vs. “before/as” (n=9), “until” vs. “after/when/at” (n=11)
Process verbs: “until” vs. “before/as” (n=15), “until” vs. “after/when/at” (n=5)

was required on-line. For example, Piñango et al. (1999, 2006) found an effect of enrichment through lexical decision times (RTs) beginning after a delay of 250 ms after enrichment was licensed; data from the current study suggest that the neural correlates of enrichment processing may already have begun indexing an enrichment process by that time (e.g., starting as late as the 100-300 ms window, perhaps while the licensing word is being processed, long before participants in Piñango et al.'s study had begun to make their responses). Indeed, these current results speak directly to the observation that Piñango et al.'s (2006) were unable to find increased processing costs at the 0 ms delay point, a result which, given the current findings, may very well be due to conflation of their modifier types.

Ultimately, the results of this study emphasize that caution must be taken when interpreting results based on heterogeneous sets of experimental stimuli. Indeed, this experiment shows that choices made regarding which temporal modifier is used to license coercion have a significant influence on whether coercion is, in fact, manifested given a particular sentence context. Previous investigations that conflated modifiers with different cognitive, psycholinguistic properties (such as nouns vs. prepositions) may have inadvertently obscured otherwise significant results. This underscores the importance of making theoretically vetted choices regarding which methodology is most appropriate for the psycholinguistic processes being investigated.

Table 3.1: The 2X2 experimental design crossing factors of Cardinality and Modifier Type used in Todorova et al. (2000). Object of the predicate is in *italics*, durative modifier is underlined.

	Durative modifier	Non-durative modifier
Singular object	Even though Howard sent <i>a large check</i> to his daughter <u>for many years</u> , she refused to accept his money	Even though Howard sent <i>a large check</i> to his daughter <u>last year</u> , she refused to accept his money
Plural object	Even though Howard sent <i>large checks</i> to his daughter <u>for many years</u> , she refused to accept his money	Even though Howard sent <i>large checks</i> to his daughter <u>last year</u> , she refused to accept his money

Table 3.2: Aspectual Coercion Effects: Omnibus Mean Amplitude ANOVAs in the 100-300, 300-500, and 500-1000 Millisecond Latency Ranges.

Source	df	100-300 ms		300-500 ms		500-1000 ms	
		F	P(HF)	F	P(HF)	F	P(HF)
Omnibus ANOVA (28 electrodes)							
Verb type (Point action vs. Process)	1, 18	0.154	.6997	1.617	.2197	2.071	.1673
Coercion (Enriched vs. Transparent)	1, 18	3.378	.0826~	3.059	.0973~	6.742	.0182*
Modifier (“For” vs. “Until”)	1, 18	11.423	.0033**	23.108	.0001**	15.991	.0008**
Verb type X Coercion	1, 18	1.985	.1759	3.169	.0919~	0.934	.3466
Verb type X Modifier	1, 18	5.325	.0331*	2.075	.1669	5.994	.0248*
Coercion X Modifier	1, 18	1.509	.2351	6.515	.0200*	11.159	.0036**
Verb type X Coercion X Modifier	1, 18	0.765	.3934	0.326	.5751	0.078	.7832
Verb type X Electrode	27, 486	0.831	.7120	1.284	.1561	0.628	.9290
Coercion X Electrode	27, 486	1.410	.2505	1.335	.1233	2.765	.1137
Modifier X Electrode	27, 486	3.350	.0838~	4.763	.0431*	1.382	.2551
Verb type X Coercion X Electrode	1, 18	0.765	.3934	1.548	.2294	0.699	.8706
Verb type X Modifier X Electrode	27, 486	1.396	.2528	0.785	.7726	1.555	.2284
Coercion X Modifier X Electrode	27, 486	0.780	.7797	2.285	.1480	2.677	.1192
Verb type X Coercion X Modifier X Electrode	27, 486	0.402	.9972	0.766	.7969	0.864	.6904

ms = milliseconds; df = degrees of freedom; HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1, **p* < 0.05, ***p* < 0.01

Table 3.3: Aspectual Coercion Effects For Verb Type: Mean ERP Amplitude ANOVAs in the 100-300, 300-500, and 500-1000 Millisecond Latency Ranges, by Modifier type (“For...” vs. *at/until*).

		100-300 ms		300-500 ms		500-1000 ms	
<i>Source</i>	<i>df</i>	<i>F</i>	<i>P(HF)</i>	<i>F</i>	<i>P(HF)</i>	<i>F</i>	<i>P(HF)</i>
<i>Omnibus ANOVA (28 electrodes)</i>							
<i>Modifier (“For” vs. “at/until”)</i>							
<i>Verb type</i>							
Point action	1, 18	17.459	.0006**	20.067	.0003**	32.126	<.001**
Process	1, 18	0.455	.5088	3.171	.0918~	0.320	.5785

ms = milliseconds; df = degrees of freedom; HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1

**p* < 0.05

***p* < 0.01

Table 3.4: Aspectual Coercion Effects by Modifier Type: Mean ERP Amplitude ANOVAs in the 100-300, 300-500, and 500-1000 Millisecond Latency Ranges.

		100-300 ms		300-500 ms		500-1000 ms	
Source	df	F	P(HF)	F	P(HF)	F	P(HF)
Omnibus ANOVA (28 electrodes)							
“For...” Type Modifier							
Verb type	1, 18	2.833	.1096	3.672	.0713~	0.951	.3423
Coercion	1, 18	3.912	.0635~	8.832	.0082**	19.798	.0003**
Verb type X Coercion	1, 18	0.370	.5505	0.626	.4392	0.235	.6337
At/Until Type Modifier							
Verb type	1, 18	3.358	.0835~	0.311	.5842	7.028	.0163*
Coercion	1, 18	0.769	.3922	0.246	.6258	0.233	.6348
Verb type X Coercion	1, 18	2.549	.1278	3.298	.0861~	0.765	.3933

ms = milliseconds; df = degrees of freedom; HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1

**p* < 0.05

***p* < 0.01

Table 3.5: Aspectual Coercion Effects by Verb Type: Mean ERP Amplitude ANOVAs in the 100-300, 300-500, and 500-1000 Millisecond Latency Ranges.

		100-300 ms		300-500 ms		500-1000 ms	
Source	df	F	P(HF)	F	P(HF)	F	P(HF)
Omnibus ANOVA (28 electrodes)							
Point action verbs							
Modifier type	1, 18	17.459	.0006**	20.067	.0003**	32.126	<.001**
Coercion	1, 18	0.356	.5582	0.045	.8350	1.366	.2578
Modifier type X Coercion	1, 18	3.386	.0823~	5.947	.0253*	10.448	.0046**
Process verbs							
Modifier type	1, 18	0.455	.5088	3.171	.0918~	0.320	.5785
Coercion	1, 18	5.771	.0273*	9.815	.0057**	7.694	.0125*
Modifier type X Coercion	1, 18	0.132	.7202	2.257	.1503	3.759	.0684~

ms = milliseconds; df = degrees of freedom; HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1

**p* < 0.05

***p* < 0.01

Table 3.6: Aspectual Coercion Effects by Verb Type Only in “For...” Type Modifiers: Mean ERP Amplitude ANOVAs in the 100-300, 300-500, and 500-1000 Millisecond Latency Ranges.

		100-300 ms		300-500 ms		500-1000 ms	
Source	df	F	P(HF)	F	P(HF)	F	P(HF)
Omnibus ANOVA (28 electrodes)							
“For...” Type Modifier, Point action verb							
Coercion	1, 18	1.875	.1898	2.901	.1057	13.141	.0019**
“For...” Type Modifier, Process verb							
Coercion	1, 18	3.664	.0716~	7.715	.0124*	8.886	.0080**

ms = milliseconds; df = degrees of freedom; HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1

**p* < 0.05

***p* < 0.01

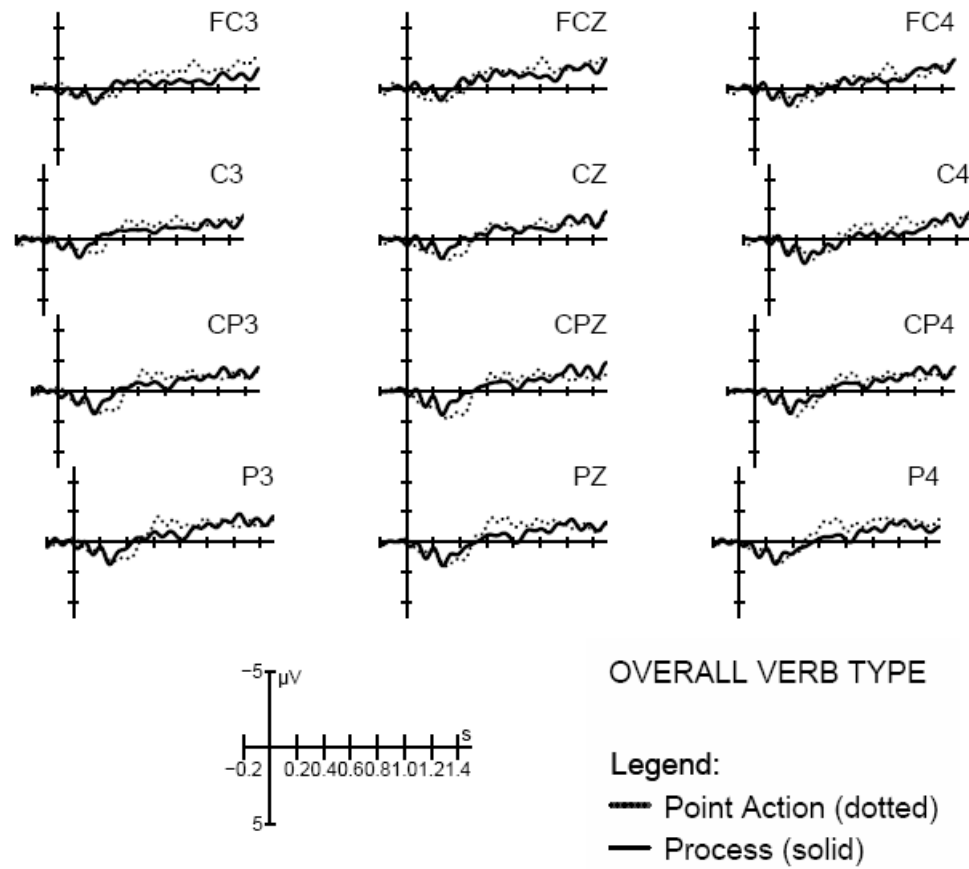


Figure 3.1: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed (adverbial or prepositional modifier), across point action (dotted) and process (solid) verbs.

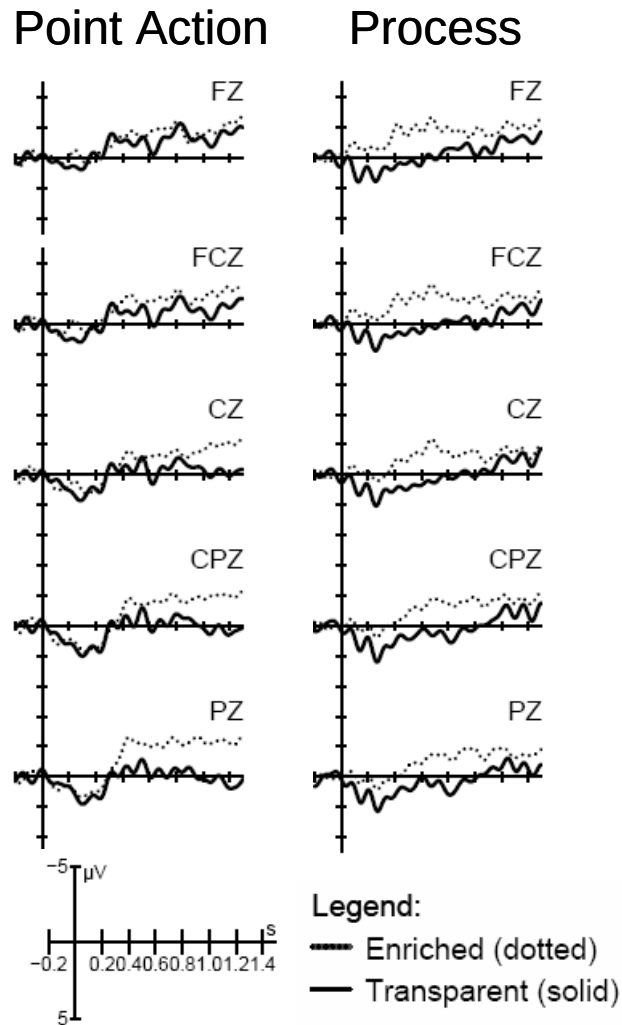


Figure 3.2: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed (adverbial or prepositional modifier), across point action (left) and process (right) verbs when in an enriching (dotted) or transparent (solid) coercion context. Frontal electrodes are shown to emphasize the different scalp distributions between the two verb types.

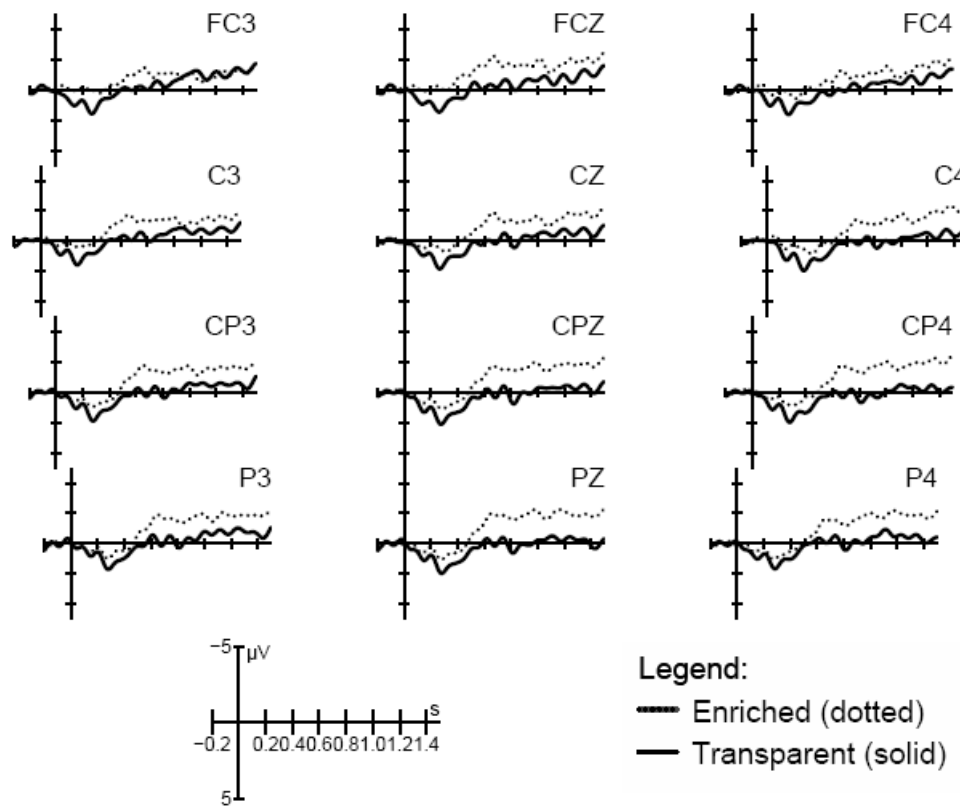


Figure 3.3: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed (adverbial or prepositional modifier), across enriched (dotted) and transparent (solid) contexts.

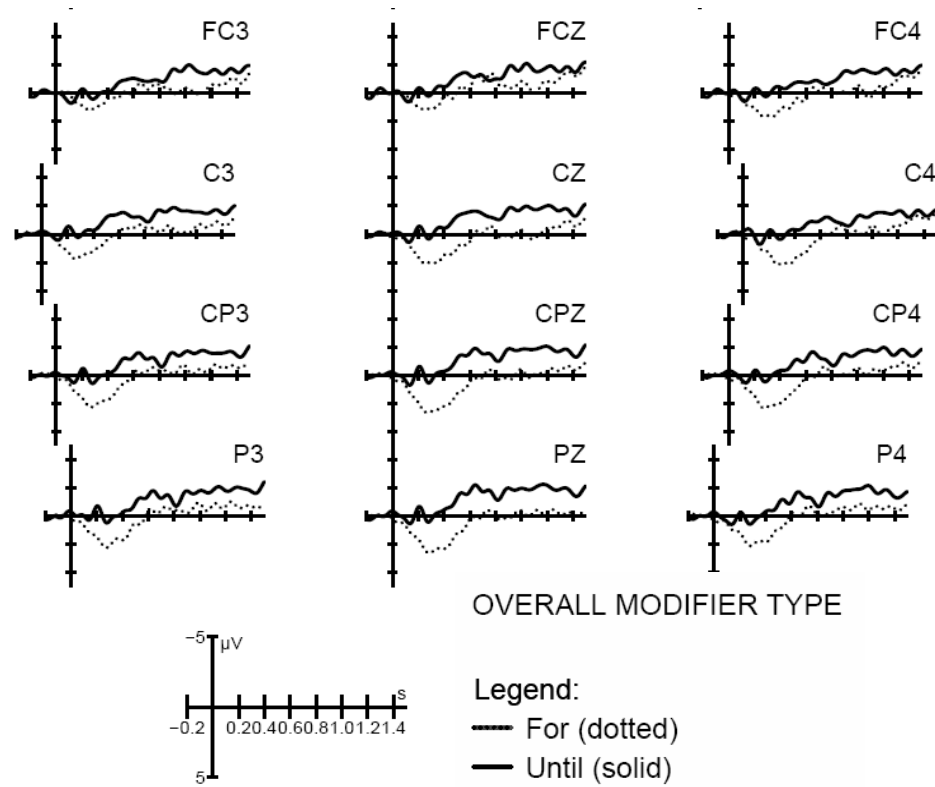


Figure 3.4: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed (adverbial or prepositional modifier), across “*For...*” (dotted) and *at/until* (solid) modifier types.

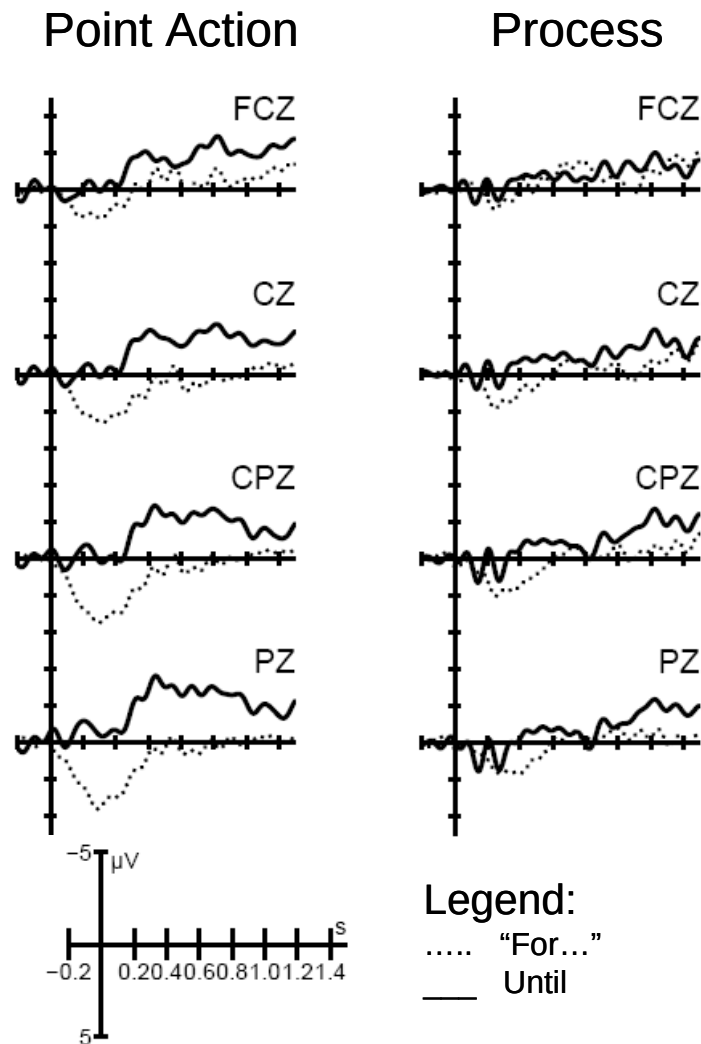


Figure 3.5: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed (adverbial or prepositional modifier). This figure displays the interaction between Verb type (Point action vs. Process) and Modifier type (“For...” vs. *at/until*) over midline electrodes.

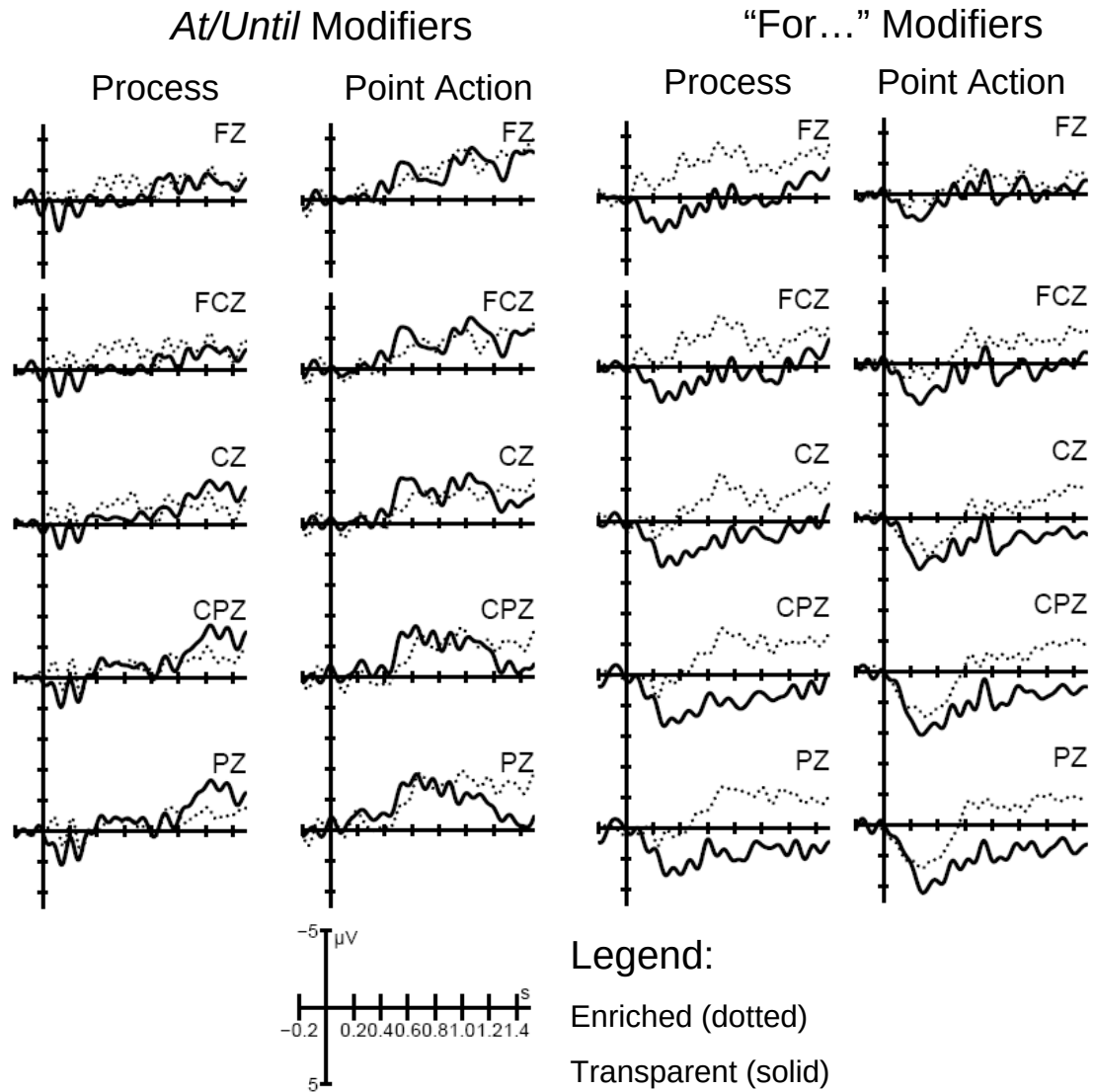


Figure 3.6: Grand average ERPs time-locked to the onset of the point where aspectual coercion is licensed, over midline electrodes. This figure displays the interaction between modifier type (*at/until* on the left and “For...” on the right), Verb type (Point action vs. Process), and coercion condition (enriched is represented by the dotted line, transparent by the solid line).

APPENDIX A

EXPERIMENTAL STIMULI

The experimental sentence pairs are divided into 4 parts. The first half of the sentences are the experimental sentences (containing point action verbs) and the second half are the control sentences (made up of process, or intrinsically continuous, verbs). The first half of each (either experimental or control) is made up of “for constructions” (for an hour vs. for an apple) and the second half are “until/after” constructions. The numbering of the sentences is consistent with the numbering of the stimuli sentences used in the study. For each pair, the intended extended interpretation is presented first.

The verb of interest is in bold, and the modifier coercing the interpretation is expressed in italics. The point at which the ERP triggers were set is marked by an asterisk (*).

Point Action

“For” Constructions

- 1 The little girl **dove** into the pool *for a long time** before deciding to go home.
The little girl **dove** into the pool *for a penny** before deciding to go home.
- 2 The eccentric foreigner **belched** after dinner *for some time** to make everyone laugh.
The eccentric foreigner **belched** after dinner *for some friends** to make everyone laugh.
- 3 The energetic personal trainer **jumped** in the gym *for a long time** after teaching a class.
The energetic personal trainer **jumped** in the gym *for a towel** after teaching a class.
- 4 The hyperactive teenager **burped** in the classroom *for a minute** after betting his friend.
The hyperactive teenager **burped** in the classroom *for a dollar** after betting his friend.
- 5 The superstitious woman **curtseyed** in the castle *for a while** as she waited for the queen.
The superstitious woman **curtseyed** in the castle *for a duke** as she waited for the queen.
- 6 The graceful ballerina **leapt** in her home *for a while** to amuse her daughter.
The graceful ballerina **leapt** in her home *for a vase** to amuse her daughter.
- 7 The tall tailor **knocked** on the door *for a minute** hoping his sister would answer.
The tall tailor **knocked** on the door *for a second** hoping his sister would answer.
- 8 The hospital worker **struck** the patient *for an hour**, after which he became remorseful.
The hospital worker **struck** the patient *for an offense**, after which he became remorseful.
- 9 The psychiatric patient **slammed** the door *for a long time** after being ignored all day.
The psychiatric patient **slammed** the door *for attention** after being ignored all day.
- 10 The television reporter **slapped** the cameraman *for minutes** after which she was fired.
The television reporter **slapped** the cameraman *for revenge** after which she was fired.
- 11 The football player **threw** the ball *for the whole day** in awe at the size of the crowd.
The football player **threw** the ball *for the award** in awe at the size of the crowd.

- 12 The busy father **tossed** the ball *for some time** to the neighbor before leaving for work.
The busy father **tossed** the ball *forcefully** to the neighbor before leaving for work.
- 13 The drunk orthodontist **slugged** the bartender *for minutes** after witnessing the bartender ogle his girlfriend.
The drunk orthodontist **slugged** the bartender *for staring** after witnessing the bartender ogle his girlfriend.
- 14 The college coed **stabbed** the mugger *for hours** before coming to her senses.
The college coed **stabbed** the mugger *for revenge** before coming to her senses.
- 15 The young boy **punched** the school bully *for a minute**, hoping he would impress the girl in the blue dress.
The young boy **punched** the school bully *for attention**, hoping he would impress the girl in the blue dress.
- 16 The tortured prisoner **knifed** the warden *for a long time** then slipped away quietly back to his cell.
The tortured prisoner **knifed** the warden *for a good friend** then slipped away quietly back to his cell.
- 17 The school librarian **poked** the sleeping boy *for minutes** before he finally woke up.
The school librarian **poked** the sleeping boy *for drooling** before he finally woke up.
- 18 The flirtatious woman **kissed** the man *for the afternoon** after meeting him only the day before.
The flirtatious woman **kissed** the man *for the help** he gave after meeting him only the day before.
- 19 The artistic teacher **cut** the paper *for a minute** for the class's crafts project.
The artistic teacher **cut** the paper *for a second** for the class's crafts project.
- 20 The stylish woman **winked** noticeably *for the afternoon**, hoping the stranger would notice her.
The stylish woman **winked** noticeably *for the gentleman**, hoping the stranger would notice her.

“Until” Constructions

- 21 The frustrated youth **banged** on the table *until** his demands were met with opposition.
The frustrated youth **banged** on the table *before** his demands were met with opposition.
- 22 The rookie policeman **shot** at the kidnapping suspect *until** the hostage got away.
The rookie policeman **shot** at the kidnapping suspect *after** the hostage got away.
- 23 The young homemaker **swatted** at the fly *until** it flew out the window into the yard.
The young homemaker **swatted** at the fly *as** it flew out the window into the yard.
- 24 The boxer **swung** at his opponent *until** the man began looking ill.
The boxer **swung** at his opponent *after** the man began looking ill.
- 25 The tennis instructor **sneezed** into the tissue *until** his sister entered the room.
The tennis instructor **sneezed** into the tissue *when** his sister entered the room.
- 26 The annoyed student **interrupted** the professor *until** it seemed as if he was coming to the end of his phone conversation.

The annoyed student **interrupted** the professor *when** it seemed as if he was coming to the end of his phone conversation.

- 27 The exhausted mother **spanked** the boy *until** he ran away into the street.
The exhausted mother **spanked** the boy *after** he ran away into the street.
- 28 The baseball player **pitched** the ball *until** his left arm began to throb.
The baseball player **pitched** the ball *when** his left arm began to throb.
- 29 The postal worker **stamped** the paper *until** he realized he had made a mistake.
The postal worker **stamped** the paper *before** he realized he had made a mistake.
- 30 The monkey **bit** the zoo worker *until** it seemed it might be able to get free.
The monkey **bit** the zoo worker *after** it seemed it might be able to get free.
- 31 The dangerous burglar **clubbed** the storeowner *until** the police arrived at the scene.
The dangerous burglar **clubbed** the storeowner *before** the police arrived at the scene.
- 32 The little boy **socked** the teacher *until** his mother calmed him down.
The little boy **socked** the teacher *before** his mother calmed him down.
- 33 The elderly doorman **yawned** *until** the tenant asked if he would help her.
The elderly doorman **yawned** *before** the tenant asked if he would help her.
- 34 The killer bee **stung** the bird *until** a flock of sea gulls attacked the beehive.
The killer bee **stung** the bird *after** a flock of sea gulls attacked the beehive.
- 35 The frustrated actress **hit** her co-star *until** the writers agreed to put her in the scene.
The frustrated actress **hit** her co-star *before** the writers agreed to put her in the scene.
- 36 The spoiled little girl **pinched** the older boy *until** he took away her new toy.
The spoiled little girl **pinched** the older boy *when** he took away her new toy.
- 37 The well dressed woman **blinked** *until** the gusty wind blew past her face.
The well dressed woman **blinked** *after** the gusty wind blew past her face.
- 38 The salesman **rang** the doorbell *until** a disheveled man answered the door.
The salesman **rang** the doorbell *before** a disheveled man answered the door.
- 39 The convict **kicked** the lawyer *until** he was promised a quick release.
The convict **kicked** the lawyer *before** he was promised a quick release.
- 40 The streetlight **flashed** *until** dawn, signaling the beginning of a new day.
The streetlight **flashed** *at** dawn, signaling the beginning of a new day.

Process verbs

“For” constructions

- 1 The bride **hugged** the photographer *for a minute** until she saw the pictures he had taken.
The bride **hugged** the photographer *for a discount** until she saw the pictures he had taken.
- 2 The magician **painted** the door *for a while** to annoy his visiting relatives.

- The magician **painted** the door *for a friend** to annoy his visiting relatives.
- 3 The investigator **followed** the woman *for a long time** before going home to have dinner.
The investigator **followed** the woman *for a client** before going home to have dinner.
- 4 The young lady **serenaded** her lover *for a whole day** in celebration of his birthday.
The young lady **serenaded** her lover *for attention** in celebration of his birthday.
- 5 The upset little boy **dragged** the dog *for hours** as he tried to find a doctor.
The upset little boy **dragged** the dog *for help** as he walked through the park to find a doctor.
- 6 The boy **rocked** the baby *for a minute** until he realized that it had fallen asleep.
The boy **rocked** the baby *for a dollar** until he realized that it had fallen asleep.
- 7 The dentist **drilled** a tooth *for a long time** stopping only to take a few sips of water.
The dentist **drilled** a tooth *for a patient** stopping only to take a few sips of water.
- 8 The athlete **swam** through the ocean *for hours** before meeting with her coach at the gym.
The athlete **swam** through the ocean *for fun** before meeting with her coach at the gym.
- 9 The man **recited** the material *for the whole day** before he was called in to make the presentation.
The man **recited** the material *for the speech** before he was called in to make the presentation.
- 10 The homeless man **meandered** in the park *for a minute** after seeing an old friend.
The homeless man **meandered** in the park *for a moment** after seeing an old friend.
- 11 The lead singer **sang** with his band *for hours** until the neighbors called the police.
The lead singer **sang** with his band *for money** until the neighbors called the police.
- 12 The skating instructor **glided** on the ice *for a while** before hurting her ankle.
The skating instructor **glided** on the ice *for a class** before hurting her ankle.
- 13 The snake **hissed** at the mouse *for a minute** until the rodent was able to escape.
The snake **hissed** at the mouse *for a second** until the rodent was able to escape.
- 14 The performer **danced** with her partner *for hours** while dreaming about becoming famous.
The performer **danced** with her partner *for practice** while dreaming about becoming famous.
- 15 The radio DJ **stared** at the woman *for a minute** during the commercial break.
The radio DJ **stared** at the woman *for a moment** during the commercial break.
- 16 The flight attendant **wept** on the captain's shoulder *for a minute** until he agreed to give her a day off.
The flight attendant **wept** on the captain's shoulder *for attention** until he agreed to give her a day off.
- 17 The man **breathed** loudly *for a long time** before his friends finally made him stop.
The man **breathed** loudly *for attention** before his friends finally made him stop.
- 18 The light **shone** *for a couple hours** after the last cab left the parking lot.
The light **shone** *for a pedestrian** after the last cab left the parking lot.

- 19 The man **whistled** loudly *for a while** until he met up with an old friend for lunch.
The man **whistled** loudly *for a crowd** until he met up with an old friend for lunch.
- 20 The secretary **typed** quietly *for the hour** before deciding to go home.
The secretary **typed** quietly *for the doctor** before deciding to go home.

“Until” constructions

- 21 The girl **slumbered** in the hall *until** her mom came to pick her up.
The girl **slumbered** in the hall *before** her mom came to pick her up.
- 22 The impatient teenager **squeezed** her friend’s hand *until** she was given all the information.
The impatient teenager **squeezed** her friend’s hand *before** she was given all the information.
- 23 The old man **slept** in the cabin *until** his friend came back with dinner.
The old man **slept** in the cabin *before** his friend came back with dinner.
- 24 The sleeping man **dreamed** about his family *until** his wife walked in and woke her up.
The sleeping man **dreamed** about his family *before** his wife walked in and woke her up.
- 25 The new student **cried** in the classroom *until** the teacher said he could go home.
The new student **cried** in the classroom *before** the teacher said he could go home.
- 26 The president **lounged** in the hall *until** his name was called to enter.
The president **lounged** in the hall *before** his name was called to enter.
- 27 The man **napped** in his bed *until** it was time for him to go to work.
The man **napped** in his bed *when** it was time for him to go to work.
- 28 The little girl **dogpaddled** in the pool *until** the teacher told her to take a break.
The little girl **dogpaddled** in the pool *before** the teacher told her to take a break.
- 29 The man **rested** on the beach *until** the tourists became too noise.
The man **rested** on the beach *before** the tourists became too noise.
- 30 The cat **chewed** on the toy mouse *until** it spotted a rat dart across the room.
The cat **chewed** on the toy mouse *before** it spotted a rat dart across the room.
- 31 The criminal **peered** at the witnesses *until** he finally gave himself up to the police.
The criminal **peered** at the witnesses *before** he finally gave himself up to the police.
- 32 The little girl **snoozed** in the hallway *until** the dog began barking loudly.
The little girl **snoozed** in the hallway *before** the dog began barking loudly.
- 33 The employee **glared** at his boss *until** he was asked to leave the building.
The employee **glared** at his boss *when** he was asked to leave the building.
- 34 The hungry puppy **ate** at the table *until** his owner redirected him to the kitchen.
The hungry puppy **ate** at the table *before** his owner redirected him to the kitchen.
- 35 The criminal **chased** the girl *until** the police came to her aid in the park.
The criminal **chased** the girl *before** the police came to her aid in the park.

- 36 The librarian **cleaned** the house *until** her husband called her for dinner.
The librarian **cleaned** the house *before** her husband called her for dinner.
- 37 The telephone **buzzed** loudly *until** noon when the butler finally picked it up.
The telephone **buzzed** loudly *at** noon when the butler finally picked it up.
- 38 The woman **caressed** the baby's cheek *until** he began crying loudly.
The woman **caressed** the baby's cheek *before** he began crying loudly.
- 39 The active man **perspired** heavily *until** someone opened the door to the house
The active man **perspired** heavily *when** someone opened the door to the house
- 40 The boy **slumped** impatiently *until** the teacher asked him to leave the room.
The boy **slumped** impatiently *when** the teacher asked him to leave the room.

REFERENCES

- Briscoe, E, Copestake, A, & Bougarev, B (1990). Enjoy the paper: Lexical semantics via lexicology. In proceedings of the 13th international conference on Computational Linguistics, Helsinki, 42-47.
- Brown, CM, Hagoort, P, & ter Keurs, M (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience*, 11, 261-281.
- Chwilla, DJ, Brown, CM, & Hagoort, P (1995). The N400 as a function of the level of processing. *Psychophysiology*, 32, 274-85.
- Clifton, C & Frazier, L (1989). Comprehending sentences with long-distance dependencies. In M. Tanenhaus & G Carlson (eds.), Linguistic structure in language processing. Dordrecht: Kluwer.
- Coulson, S (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. London: Cambridge University Press.
- De Swart, H (1998). Aspect shift and coercion. *Natural Language and Linguistic Theory*, 16, 347-385.
- Jackendoff, R (1997). *The architecture of the language faculty*. *Linguistic inquiry monograph*. Cambridge, MA: MIT Press.
- King, JW, & Kutas, M (1995). A brain potential whose latency indexes the length and frequency of words. *Center for Research in Language Newsletter*, 10-2.
- McElree, B, Frisson, S, & Pickering, MJ (2006). Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. *Cognitive Science*, 30, 181-192.
- McElree, B, & Griffith, T (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 134-157.
- McElree, B, Pylkkänen, L, Pickering, MJ, & Traxler, MJ (2006). The time course of enriched composition. *Psychonomic Bulletin and Review*, 13, 53-59.
- McElree, B, Traxler, MJ, Pickering, MJ, Seely, R, & Jackendoff, R (2001). Reading time evidence for enriched composition. *Cognition*, 78, B17-B25.
- Mourelatos, A (1981). Events, processes, and states. *Linguistics and Philosophy*, 2, 415-434.

- Münste, TF, Wieringa, BM, Weyerts, H, Szentkuti, A, Matzke, M, & Johannes, S (2001). Differences in brain potentials to open and closed class words: Class and frequency effects. *Neuropsychologia*, 39, 91-102.
- Pickering, MJ, McElree, B, Frisson, S, Chen, L, & Traxler, MJ (2006). Underspecification and aspectual coercion. *Discourse Processes*, 42, 131-155.
- Piñango, MM, Winnick, A, Ullah, R, & Zurif, E (2006). Time-course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research*, 35, 233-244.
- Piñango, MM, Zurif, E, & Jackendoff, R (1999). Real-time processing implications of aspectual coercion at the syntax-semantics interface. *Journal of Psycholinguistic Research*, 28, 395-414.
- Proctor, AS, Dickey, MW, & Rips, LJ (2004 August). *The time-course and cost of telicity inferences*. Proceeding of the 26th Annual Meeting of the Cognitive Science Society, Chicago, IL.
- Pustejovsky, J (1991). The generative lexicon. *Computational Linguistics*, 17, 409-441.
- Pustejovsky, J (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Todorova, M, Straub, K, Badecker, W, & Frank, R (2000a). Aspectual coercion and online computation of sentential aspect. Cognitive Science Society Annual Meeting, Philadelphia, PA.
- Todorova, M, Straub, K, Badecker, W, & Frank, R (2000b). Processing correlates of aspectual computation. Presented to the Workshop on Events and Paths, ESSLI XII, Birmingham, England.
- Traxler, MJ, McElree, B, Williams, RS, & Pickering, MJ (2005). Context effects in coercion: Evidence from eye-movements. *Journal of Memory and Language*, 53, 1-25.
- Traxler, MJ, Pickering, MJ, & McElree, B (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language*, 4, 530-547.

Chapter 4

Abstract

Previous studies have shown the parser to be sensitive to verb subcategorization violations, indicating that this syntactic lexical property is available and utilized almost immediately on-line. What has not been demonstrated conclusively, however, is whether a verb's subcategorization frames are independent of the semantics of the nouns that fill them. This study employs event-related potentials (ERPs) time-locked to violations in English sentences and sentences mostly devoid of semantic information – so-called “Jabberwocky” sentences – to test this question. Results from this study find that the early ERP response associated with phrase structure building (Left Anterior Negativity, or LAN) that has been found previously in English and Jabberwocky sentences was absent in grammatical violations in both these languages, while the late “syntactic repair” ERP response (P600) was elicited by violations in both English and Jabberwocky sentences. These results suggest that semantic content is not necessary for late phrase-structure recovery mechanisms, as has been argued previously.

4.1. INTRODUCTION

Although most psycholinguistic models of language comprehension are aligned with regard to the distinction of semantic and syntactic linguistic information, there exists some divergence among the models as to the hypothesized contributions of semantic and syntactic lexical properties to successful processing, as well as to the relative time course of the informational availability/use. So called “syntax first” models, for example, posit that the initial stage of sentence processing is primarily concerned with aligning syntactic properties of the linguistic components; that is, a sentence structure is built based on syntax first, independent of the conceptual meaning, which only becomes relevant at a later stage (Frazier, 1987; Friederici, 1995, 2002). For example, such models predict that the sentence (1a) would be processed grammatically, even though the nouns have no semantic content (i.e., they are pseudowords). This is because the initial verb is construed as being in a relative clause (i.e., signaled by “that the...”), and thus a matrix verb is necessary; “staged” satisfies this requirement. On the other hand, sentence (1b) lacks a matrix verb and would be found to be ungrammatical.

(1a) The lufnuts that the easmet hated staged loud sebots after the pob.

(1b) The lufnuts that the easmet hated loud sabots after the pob.

Thus, syntax-first models would predict that the syntactic computation is performed with each word, and grammaticality (or, in this case, ungrammaticality) would be processed immediately (e.g., Clifton & Frazier, 1989). The current study tests this prediction by using both English sentences and sentences which are semantically

impoverished, sometimes called “Jabberwocky” sentences (Canseco-Gonzalez, et al., 1997; Hahne & Jescheniak, 2001).

The event-related potentials methodology (ERPs) has proven reliable in indexing linguistic processing such as that involved in this investigation. Three particular ERP components have been delineated historically in reference to language processing, associated with either semantic or syntactic information (or, as mentioned below, with both). Following a brief review of the ERP components typically involved in language studies, this paper describes a study – part of a larger set of studies – designed to examine the relative time-course and contribution of syntactic and semantic information to the on-line processing of sentences in the auditory modality.

4.1.1. Language related ERP effects

4.1.1.1. Semantic processing

A component widely studied in psycholinguistic inquiries is the N400, found by Kutas and Hillyard (1980) in response to semantic violations such as “socks” in the sentence “I like my coffee with sugar and socks.” The N400 effect is described as greater negative amplitude in response to items which are semantic violations as compared to non-violations, having its peak ~400 ms after the violation is presented, and having a centroparietal maximum. In visual studies it is often greater over the right than over the left, while in auditory studies it tends to be more bilaterally distributed. The N400 was originally characterized as a violation of semantic

expectancy (Kutas & Hillyard, 1980, 1983; Kutas, 1983); further studies have suggested that it may be a more general index of difficulty integrating an item into an ongoing semantic context and is often seen to be a function of an item's cloze probability (Van Petten & Kutas, 1991; Kutas & Van Petten, 1984). Recently, Hagoort, Hald, Bastiaansen, and Petersson (2004) showed that the N400 may also index difficulties in the on-line evaluation and integration of real-world knowledge.

4.1.1.2. Syntactic processing

Two components have been elaborated which appear to be related – at least primarily – to syntactic processing. The first of these is a left anterior negativity (LAN), which is typically seen in the same time window as the N400 (300-500 ms) but with a left, anterior maximum. An early LAN has been found in response to phrase structure violations (Friederici, Pfeifer, & Hahne, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991), while the LAN has been elicited by subject-verb number agreement violations (Coulson, King, & Kutas, 1998; Friederici & Mecklinger, 1996; Hagoort & Brown, 2000; Osterhout & Mobley, 1995; De Vincenzi et al., 2003), violations of phrase structure (Hagoort & Brown, 2000; Brown, Lehmann, & Marsh, 1980; Osterhout & Holcomb, 1992; Neville, Nicol, Barss, Forster, & Garrett, 1991;), and long-distance constructions involving, for example, gap filling or wh-movement phenomena (Kluender & Kutas, 1993; King & Kutas, 1995; Fiebach, Schleewsky, & Friederici, 2002). An early version of the LAN, referred to as the ELAN by some researchers, has been elicited following word category violations (Neville et al., 1991;

Hahne & Friederici, 1999; Friederici, Hahne, & Mecklinger, 1996), and has been taken by some to be an index of the earliest stage of phrase structure building, in which syntactic category information is aligned

A later component, the P600 (also known as a syntactic positive shift, or late positive complex), is a positivity in the ERP, typically having a maximal amplitude between 500 and 800 ms post-presentation and having a centroparital distribution. Although it has been elicited by precisely the types of violations described above, the P600 is not necessarily viewed as indexing syntactic violations. Indeed, it has been elicited by “garden path” sentences (Osterhout & Holcomb, 1992; Osterhout, Holcomb, & Swinney, 1994; McKinnon & Osterhout, 1996; Friederici & Mecklinger, 1996) syntactically ambiguous sentences (Osterhout & Holcomb, 1992; Friederici, 1995), and sentences with increased syntactic complexity that are nevertheless well-formed (Fiebach, Schlesewsky, & Friederici, 2002; Kaan et al., 2000; Phillips, Kazanina, Wong, & Ellis, 2001; Featherston, Gross, Münte, & Clahsen, 2000). The P600 has also recently been suggested to index general reprocessing as a monitoring or error-checking mechanism (van Herten, Chwilla, & Kolk, 2006; van Herten et al., 2005; Kolk et al., 2003).

4.1.1.3. Combined violations

The clean dissociation between linguistic processes and their corresponding ERP components is not perfect, however. Numerous studies have employed conditions which include violations of semantics, violations of syntax, and combined

violations (e.g., the same word constitutes a semantic and syntactic violation) to attempt to validate the linguistic specificity of these components, with various outcomes (e.g., Rösler, Friederici, Putz, & Hahne, 1993; Ainsworth-Darnell, Shulman, & Boland, 1998; Gunter, Friederici, & Shriefers, 2000; Gunter, Stowe, & Mulder, 1997; Hahne & Friederici, 1998, 2001; Osterhout & Nicol, 1999; to name a few). In some cases of combined violation, stimuli have elicited a (E)LAN followed by a P600 in the absence of an N400, suggesting that if initial parsing attempts encounter a breakdown, then no attempt at semantic analysis is attempted (Hahne & Jescheniak, 2001). Other studies find that semantic and syntactic violations are processed together, resulting in both an N400 and a P600. For example, Osterhout and Nicol (1999) used stimuli of the form “One kangaroo at the San Diego Zoo would sometimes sit / sitting / write / writing all day”, and the double violation (“writing”) elicited both an N400 and a P600. Interestingly, the purely semantic violation “sitting” elicited a LAN in one experiment and an N400 in another (though this may have simply been due to electrode grouping choices).

4.1.2. Thematic assignment issues

Of particular note for the current investigation are results in which a P600 has been elicited by lexico-semantic violations in syntactically well-formed sentences in the absence of an N400 (Kim & Osterhout, 2005; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, in press). For example, Kuperberg and

colleagues (2003, 2006) have shown that a P600 is elicited by verbs such as “eat” in sentences such as “At breakfast, the eggs would eat every day.” The N400, which would be expected in such cases of violation of semantic and pragmatic expectancy, was not found. Kuperberg, Sitnikova, et al. (2003) have suggested that the P600 may be the result of a thematic mismatch between the verb and its preceding argument(s). These authors suggest that the verb “eat” assigns the thematic role of Agent to an animate argument; in the case of an inanimate argument (“the eggs”), the role of Theme would be more likely. In cases like this, the semantic context of the verb and its arguments are preserved; that is, eggs are conceptually associated with the process of eating, and because the parser experiences no difficulty in conceptually matching “eat” with “eggs”, such a violation does not generate an N400 (Kim & Osterhout, 2004; Kuperberg, Sitnikova, et al., 2003; Kuperberg, Kerher, et al., 2006).

4.1.3. “Jabberwocky”

A handful of studies have attempted to get around potential semantic/thematic effects in teasing apart the functional significance of the (E)LAN and P600. Canseco-Gonzalez et al. (1997; in prep) studied visually presented English verbs placed in English sentences and sentences with pseudo-word (“Jabberwocky” words) content words in an attempt to eliminate the verb argument semantic effects. Across two violation types (phrase structure and verb subcategorization, which the authors considered a type of phrase structure violation), Canseco-Gonzalez and colleagues found a LAN in response to violations in both English and Jabberwocky sentences, but

a P600 effect to violations in only the English sentences. This suggested to the authors that the linguistic processes underlying the LAN operate automatically, independently of the semantic content of the sentence, while the processes underlying the P600 require semantic content. In other words, the parser only attempts to reanalyze or reconstruct sentences which have some meaningful content.

Münter, Matzke, and Johannes (1997) studied verb number agreement errors in visually presented German and pseudo-word sentences. These authors found an early negativity (LAN) in violations of morphosyntax in German but not in pseudo-word sentences. They also found a P600 in response to German sentences with the violation, but no such positivity for pseudo-word sentences, suggesting that the P600 indexes morphosyntactic violations provided there is sufficient semantic content in the stimulus.

More recently, Hahne and Jescheniak (2001) conducted an auditory ERP study to examine this dissociation using phrase structure violations, which have been shown to produce an ELAN followed by a P600 (Friederici, Pfeifer, & Hahne, 1993; Gunter, Friederici, & Hahne, 1999; Hahne & Friederici, 1999). They included German and Jabberwocky sentences which were either syntactically correct or incorrect. These authors found an ELAN followed by a P600 in both the German and Jabberwocky sentences. Finding a P600 following a phrase structure violation in Jabberwocky appears to contradict the results of Canseco-Gonzalez et al. (1997) and Münter et al. (1997), both of which failed to find P600 following semantically empty syntactic violations.

Although Canseco-Gonzalez et al. (1997) presented phrase structure violations designed to elicit the same kind of processing as the sentences in Hahne & Jescheniak (2001), the presence of the P600 differed between the studies. One possible explanation for the divergent results may lie in the presentation modality across these studies. In an attempt to elicit the early syntactic processing component (ELAN), Canseco-Gonzalez et al. used visual stimuli while Hahne & Jescheniak used auditory. It may be that the different results are due to different processing demands or strategies based on whether comprehenders engage in auditory vs. visual processing. Although some studies find no difference in the evoked brain responses to visual vs. auditory stimuli (e.g., Balconi & Pozzoli, 2005), there is reason to suspect that task-related processing differences may contribute to differences in overall distribution of language-related ERP components, including the presence or absence of expected effects (Connolly, Phillips, Stewart, & Brake, 1992; Friederici, Pfeifer, & Hahne, 1992; Holcomb & Neville, 1991; Osterhout & Holcomb, 1993; Woodward, Ford, & Hammett, 1993; Müller, King, & Kutas, 1997).

To address these possibilities, the following study was undertaken. Employing the identical stimuli as Canseco-Gonzalez et al. (1997) but in the auditory modality, this study attempts to distinguish to what extent the processes underlying the (E)LAN and the P600 are dependent upon semantic information.

4.1.4. Current investigation

The precise relationship between the (E)LAN and the P600 regarding semantic processing is not entirely clear. The (E)LAN may be a reliable index of the parser's attempts at processing violations of phrase structure; the P600 appears to be an index of structural reanalysis or recovery of meaning following, among other things, violations of syntax (including phrase structure violations). Therefore, sentences that contain violations of phrase structure should elicit a combined (E)LAN-P600 (Neville et al., 1991; Münte, Heinze, & Mangun, 1993; Friederici, 1995; 2002; Hahne & Jescheniak, 2001). This pattern of results should be the same independent of semantic content of the stimulus (e.g., in both English and Jabberwocky sentences); regardless of whether the words have semantic content, as long as the items are perceived to be of an incorrect syntactic category, the appropriate syntactic reflex should operate (cf. Hahne & Jescheniak, 2001; Münte, Matzke, & Johannes, 1997). Because there is no explicit semantic violation in cases such as these, no N400 is predicted in either English or Jabberwocky sentences. Phrase structure I (PS1) stimuli were constructed to test this prediction.

Likewise, because the LAN has been shown to index violations of verb subcategorization (Osterhout & Holcomb, 1992; Rösler et al., 1993), a LAN-P600 complex should be present in cases in which verb subcategorization rules are violated. Results from a study by Canseco-Gonzalez and colleagues (1997) suggest that verb subcategorization violations may be processed in the same manner as phrase structure violations. In line with those results, verb subcategorization violations in the current study are considered another type of phrase structure violation (PS2). Even though

verb subcategorization may involve semantic processing in the form of thematic assignment, violations of verb subcategorization in the absence of outright semantic violations have been shown to result in a structural reprocessing cost (indexed by only the LAN-P600) in the absence of the “semantic” component, the N400; indeed, violations of thematic assignment appear to elicit a P600 and not an N400 (e.g., Kuperberg et al., 2003, 2006). This prediction was tested via the phrase structure II (PS2) stimuli.

4.2. METHOD

4.2.1. Participants

Thirty-three (33) undergraduates at the University of California, San Diego took part in this study and received course credit for their participation. All had normal or corrected to normal hearing and vision, with no history of significant neurological impairment. All were right handed (by self report on an expanded Edinburg handedness inventory (Oldfield, 1971). All participants provided informed consent. Because of excessive data rejection due to incorrect responses or movement/eyeblick artifact (see below), 9 participants were excluded from data analysis. This left 24 participants (14 were female) between the ages of 18 and 22 years (mean = 19.75 years, SD = 1.29 years).

4.2.2. Materials

Stimuli were designed to elicit phrase structure violations, the two types of which have previously produced (E)LAN-P600 effects (Canseco-Gonzalez, 1997; Neville et al., 1991; Osterhout & Holcomb, 1992).

For each violation type, there were 4 conditions, crossing language (English vs. Jabberwocky) with grammaticality (grammatical vs. ungrammatical).²⁸

Thus, for Phrase Structure 1 (PS1) violations, the sentences were of the following form. Ungrammatical sentences were produced by removing the matrix verb of the main clause. The point at which the sentence includes a violation (and the point at which the ERP was time-locked) is indicated by a caret (^):

English:

- (a) The players that the coach hated staged ^loud riots after the game.
- (b) *The players that the coach hated ^loud riots after the game.

Jabberwocky:

- (c) The lufnuts that the easmet hated staged ^loud sebots after the pob.
- (d) *The lufnuts that the easmet hated ^loud sebots after the pob.

Phrase Structure II (PS2) sentences, containing violations of verb subcategorization rules, were constructed as follows. Grammatical versions of the sentences were constructed so that the first verb encountered is, in fact, a reduced

²⁸ These materials have been pre-tested extensively, using both a paper and pencil task (pre-test #1) and a word-by-word reading stop-making-sense task (pre-test #2), which showed increased reading times at the point of violation. Details of pre-testing methods can be found in Canseco-Gonzalez, Love, Walenski, Ahrens, Swinney, & Neville (*in prep*).

relative, which is grammatical in the case where it is an embedded clause; this becomes clear at the verb “was”. Ungrammatical sentences were produced such that the initial verb can not be a reduced relative; thus, when the verb “was” is encountered, it is disallowed by the subcategorization rules of “planned”. In other words, the (a) version is grammatical because a broker can be persuaded to conceal something, whereas the (b) version is ungrammatical; a broker can not be planned to conceal anything.

English:

- (a) The broker persuaded to conceal the transaction ^was sent to jail.
- (b) *The broker planned to conceal the transaction ^was sent to jail.

Jabberwocky:

- (c) The broder persuaded to conceal the tempkishin ^was sent to julp.
- (d) *The broder planned to conceal the tempkishin ^was sent to julp.

Thirty-six (36) sentence frames were generated for each PS type, with four sentence types in each frame (a-d above; EG, EU, JG, JU, respectively), for a total of 288 experimental sentences (36 frames X 4 sentence types X 2 PS types). Sentences were pseudo-randomized such that no more than 3 of the same condition occurred in a row. The sentences were then split into four blocks such that each block had 1 version of a given sentence; for example, sentence 1a might be heard in the first block, while sentence 1d might be heard in the second block, sentence 1c in the third block, and

sentence 1b in the fourth block. Because of the randomization technique, the same sentence ID would occur in a different point in each block.

Experimental sentences were digitized on a computer (CoolEdit Pro v1.2, Syntrillium Software) by a female native English speaker with a neutral accent. The speaker was instructed and coached on how to record the stimuli in as natural a fashion as possible to reduce prosodic cues in the violation sentences (e.g., to speak naturally over the violation points) and to make the Jabberwocky sentences understandable despite their pseudoword content. Each item began with “Ready”, followed by 500 ms pause, followed by the sentence. After the sentence, a random pause occurred (750, 1000, or 1250 ms), followed by a beep which triggered participants to respond. After a 2500 ms response window, there was a 1000 ms pause, and then the next item began.

Experimental sentences (288 total) ranged from 10-16 words in length (mean = 11.4 words, SD = 0.85). Syllables per sentence ranged from 12-25 (mean = 17.4 syllables, SD = 2.54). The mean speaking rate during sentences was 5.0 syllables/second (SD = 0.57), consistent with normal speaking rates. Mean sound file duration ranged from 7093 to 9608 ms (mean = 8142 ms, SD = 483). Sentence duration ranged from 2389 to 4957 ms (mean = 3493 ms, SD = 459). Practice/filler sentences (23 total) ranged from 10-13 words in length (mean = 11.2 words, SD = 0.89). Number of syllables per sentence ranged from 14-25 (mean = 17.7 syllables, SD = 3.04). The mean speech rate was 4.9 syllables/second (SD = 0.53). Mean sound

file duration was 8315 ms (SD = 518); mean sentence duration was 3632 ms (SD = 499).

4.2.3. Procedure

Participants were seated 1.65 m in front of an LCD monitor with stereo speakers on either side. Participants were told that they were going to hear sentences that were either grammatical or not, and they were given several examples, including ungrammatical versions of both English and Jabberwocky sentences. Participants were told to listen to the sentences and to decide whether each sentence was grammatical or ungrammatical, pressing the right button with their right hand for grammatical and the left button with their left hand for ungrammatical sentences. It was emphasized that, for each response, they should ignore the fact that some of the words might be unknown to them and focus on the structure of each sentence in order to decide if the sentence was grammatical or not. They were instructed to fixate their eyes on a red dot on the LCD and to avoid moving their bodies or their eyes, or blinking, while listening to the sentences. Following the response tone, participants were allowed to relax fixation and make eye blinks and other small movements. The entire study was conducted in one 2 hour session with a total ERP recording time of 1 hour.

The experiment began with 10 practice sentences, followed by a break to provide feedback to participants. Each block was then presented, beginning with several filler sentences (4 in the first block, 3 in each consecutive block) to allow

participants to get up to speed on the task. Practice/filler sentences were of the same format as the experimental stimuli: of the 23 total practice/fillers, 14 were grammatical (8 English and 6 Jabberwocky) and 9 were ungrammatical (5 English and 4 Jabberwocky); format of the sentences were 12 PS1 and 11 PS2. Stimuli were organized in 2 lists with a counterbalanced order of presentation to control for order effects: List 1 = ABCD, List 2 = CDAB. Participants were randomly assigned to either list.

4.2.3.1. ERP procedure

An electrode cap (QuikCap, CompuMedics, Inc.) embedded with 32 sintered Ag/AgCl electrodes was applied to the scalp, with all electrodes referenced to the left mastoid. The sites involved a modified International 10-20 System, including midline (Fz, FCz, Cz, CPz, Pz, and Oz) and lateral sites (FP1/2, F3/4, F7/8, C3/4, T7/8, P3/4, P7/8, and O1/2), as well as at FC3/4, FT7/8, CP3/4, and TP7/8. In addition, bipolar loose electrodes were applied to the outer canthus of the both eyes and above and below the left eye to monitor vertical and horizontal eye movements, respectively.

The EEG signal was amplified 100 X with a SynAmps amplifier (3dB cutoff, bandpass filtered on-line between 0.15 and 30 Hz), digitized on-line at 500 Hz. Average ERPs were formed off-line from trials free of EOG and movement artifact ($>\pm 100\mu\text{V}$). As mentioned in “Participants” above, 9 participants exhibited excessive amounts of artifact (~50% of epochs were unusable due to movement and/or eye blinks) or excessively high error rates (<60% correct on grammaticality judgments on any condition; see below), and their data were excluded from later analysis. The

remaining 24 participants showed a low rate of lost trials due to artifacts (<10% across conditions).

4.2.3.2. Data Analysis

For each participant, only trials scored as correct (i.e., grammatical scored as grammatical, ungrammatical scored as ungrammatical) were included in the analyses. ERP data to the conditions of interest (2 PS types X 2 languages X 2 grammaticality conditions) were quantified by measuring the mean amplitude (in μV) separately within three latency windows determined to be, *a priori*, most appropriate for investigating language effects: 100-300 ms, 300-500 ms, and 500-1000 ms. Omnibus repeated measures ANOVAs were computed on grand averages for each time window separately. ERP measures to each violation type were then subjected to repeated measures ANOVA on condition (2 levels: violation vs. control), language (2 levels: English vs. Jabberwocky), hemisphere (2 levels: left vs. right), and anteriority (3 levels: anterior (Fz, F3/4, F7/8, FCz, FT7/8, FC3/4), central (T7/8, C3/4, Cz), and posterior (TP7/8, CP3/4, CPz, P7/8, P3/4, Pz). The Huynh-Feldt correction was imposed whenever the number of degrees of freedom in the numerator was greater than 1; all reported *p* values are corrected.

4.3. RESULTS

4.3.1. Behavioral results

Percent correct for each of the conditions is summarized in Table 1 (below).

As can be seen, accuracy on all trials was roughly comparable overall, with mean accuracy in all conditions ~81% correct.

4.3.2. ERP results

Results reported here are based on analyses of the unfiltered data. Trials in which participants responded incorrectly were discarded, as their incorrect responses made it unclear whether they processed the sentences appropriately. Data in figures has been bandpass filtered at 0.05-9.0 Hz for visualization purposes.

An initial omnibus ANOVA was carried out in each of the windows to explore effects of the two Phrase Structure violation types. Results of the rmANOVA with violation type (PS1 vs. PS2), language (English vs. Jabberwocky), and grammaticality (grammatical vs. ungrammatical), hemisphere (left vs. right), anteriority (anterior vs. central vs. posterior) found no main effect of violation type in 100-300 ms window [$F(1, 23) = 0.255$], but a main effect in both the 300-500 ms [$F(1, 23) = 9.508, p = 0.0052$] and 500-1000 ms [$F(1, 23) = 11.415, p = 0.0026$] windows.

No other main effects were found in the 100-300 ms or 300-500 ms window, but a main effect of grammaticality was found in the 500-1000 ms window [$F(1, 23) = 5.084, p = 0.0340$], such that ungrammatical items were overall more positive than grammatical items in this window (1.756 μV vs. 1.101 μV , respectively).

In addition, numerous interactions were found to be reliable. In the 100-300 ms window, the following interactions were found: phrase structure X language [$F(1,$

23) = 7.403, $p = 0.0122$]; phrase structure X anteriority [$F(2, 46) = 13.105$, $p = 0.0014$]; laterality X anteriority [$F(2, 46) = 7.084$, $p = 0.0139$]; phrase structure X language X grammaticality [$F(2, 46) = 5.996$, $p = 0.0224$]; and phrase structure X language X anteriority [$F(2, 46) = 4.584$, $p = 0.0431$]. In the 300-500 ms window, the following interactions were significant: laterality X anteriority [$F(2, 46) = 6.496$, $p = 0.0179$]; and phrase structure X grammaticality X anteriority [$F(2, 46) = 3.694$, $p = 0.0671$]. In the 500-1000 ms window, significant interactions were found between language X anteriority [$F(2, 46) = 7.598$, $p = 0.0112$]; grammaticality X anteriority [$F(2, 46) = 7.092$, $p = 0.0139$]; laterality X anteriority [$F(2, 46) = 3.603$, $p = 0.0703$]; and phrase structure X grammaticality X laterality X anteriority [$F(2, 46) = 3.434$, $p = 0.0767$].

Because there appeared to be an overall difference in ERP patterns between the two phrase structure conditions, ANOVAs were carried out separately for each phrase structure condition in each time window. These results are summarized in Tables 2-4, and elaborated below.

4.3.2.1. Phrase Structure I

ERPs elicited by phrase structure I sentences (both grammatical and ungrammatical) for both languages are shown in Figure 1.

100-300 ms (ELAN)

In phrase structure I sentences, there was an overall effect of anteriority [$F(2, 46) = 9.189$, $p = 0.0059$], with posterior electrodes more positive (1.352 μV) than

central and anterior (1.087 μV and 0.684 μV , respectively) electrodes. An interaction between language and grammaticality trended toward significance [$F(1, 23) = 3.563$, $p = 0.0718$], such that in English, grammatical and ungrammatical sentences produced similar amplitudes (0.890 and 0.820 μV , respectively), but grammatical and ungrammatical sentences in Jabberwocky produced different amplitudes, with ungrammatical sentences producing considerably more positive amplitudes overall (1.630 μV) than grammatical sentences (0.823 μV). This interaction was due to the pervasive positive amplitudes to ungrammatical Jabberwocky sentences across the head in this time window (see Figure 2), suggesting that ungrammatical Jabberwocky sentences were processed differently than ungrammatical English sentences even in the earliest time window. Finally, a laterality X anteriority interaction approached significance [$F(2, 46) = 3.770$, $p = 0.0645$], revealing that posterior electrodes were more positive over the left hemisphere than the right, while central electrodes were more positive over the right than the left hemisphere (see Figure 3).

300-500 ms (N400/LAN)

In this time window, a borderline main effect of language [$F(1, 23) = 3.591$, $p = 0.0707$] and a borderline interaction between language X laterality [$F(1, 23) = 3.552$, $p = 0.0722$] were the only effects noted. The main effect of language manifested as an overall greater positivity for Jabberwocky sentences (1.581 μV) than for English sentences (0.924 μV). The language X laterality interaction was due to a greater positivity over the right hemisphere (1.086 μV) than the left (0.762 μV) for English sentences and a greater positivity over the left hemisphere (1.681 μV) than the

right (1.481 μV) for Jabberwocky sentences. No other effects were significant in this time window for phrase structure I violations.

500-1000 ms (P600)

Two comparisons approached significance in this time window: language X anteriority [$F(2, 46) = 3.804, p = 0.0634$], and laterality X anteriority [$F(2, 46) = 4.043, p = 0.0562$]. The language X anteriority interaction revealed increasing amplitude from the anterior to central to posterior scalp regions for English sentences (0.681, 0.822, and 1.079 μV , respectively) while amplitudes decreased from the front to the back of the head for Jabberwocky sentences (1.215, 1.121, and 0.954 μV , respectively). The laterality X anteriority interaction revealed significantly greater positivity over the left hemisphere (1.313 μV) than the right (0.583 μV) in the anterior scalp region, but no differences between left and right hemispheres in central (1.033 and 0.910 μV) or posterior (1.013 and 1.020 μV) regions (see Figure 4).

4.3.2.2. Phrase Structure II

ERPs elicited by phrase structure II sentences (both grammatical and ungrammatical) for both languages are shown in Figure 5.

100-300 ms (ELAN)

In this window, a main effect of language was found [$F(1, 23) = 5.202, p = 0.0322$], with ERP responses to English sentences overall more positive (1.464 μV) than those to Jabberwocky sentences (0.817 μV). A main effect of laterality was also found, with right hemisphere electrodes overall showing greater positivity (1.302 μV)

than left (0.980 μV). A trend toward a significant laterality X anteriority interaction [$F(2, 46) = 3.770, p = 0.0690$] revealed greater positivity in electrodes over the right hemisphere than over the left hemisphere in anterior (1.412 vs. 1.055 μV) and central (1.405 vs. 0.844 μV) regions, but no difference between the hemispheres in posterior electrodes (1.088 vs. 1.040 μV , respectively). Likewise, a language X grammaticality X anteriority interaction was significant [$F(2, 46) = 3.965, p = 0.0585$], with grammatical Jabberwocky sentences more positive than ungrammatical Jabberwocky sentences at all 3 anteriority regions, while ungrammatical English sentences revealed more positive amplitudes than grammatical sentences only at posterior electrodes (see Figure 6).

To explore the main effect of language further, additional ANOVAs were conducted on each language separately, using grammaticality, laterality, and anteriority as repeated measures factors.

English

An interaction between grammaticality X anteriority [$F(2, 46) = 4.935, p = 0.0364$] was found, with responses to ungrammatical sentences being more positive (1.581 μV) than grammatical sentences (1.003 μV) over posterior electrodes, with no difference between grammatical and ungrammatical sentences over anterior (1.796 vs. 1.469 μV) or central (1.460 vs. 1.474 μV) electrodes (see Figure 7). No other main effects or interactions were found.

Jabberwocky

An interaction between laterality X anteriority was found [$F(2, 46) = 4.718, p = 0.0404$], with amplitudes over the right hemisphere considerably more positive than over the left hemisphere over anterior (1.108 vs. 0.562 μV) and central (1.111 vs. 0.451 μV) regions but not over posterior electrodes (0.860 vs. 0.811 μV). No other effects were found.

300-500 (N400/LAN)

A main effect of laterality was found [$F(1, 23) = 5.370, p = 0.0297$], with right hemisphere electrodes revealing greater positivity (2.162 μV) than left hemisphere electrodes (1.658 μV). A borderline grammaticality X anteriority interaction [$F(2, 46) = 3.201, p = 0.0868$] revealed less positive responses to ungrammatical sentences (1.708 μV) than to grammatical sentences (2.460 μV) over anterior and central (1.721 vs. 2.098 μV , respectively) electrodes (see Figure 8). An interaction between laterality X anteriority was significant [$F(2, 46) = 4.563, p = 0.0435$], with amplitudes being more positive over right hemisphere sites than over left hemisphere sites for anterior (2.365 vs. 1.812 μV) and central (2.336 vs. 1.483 μV) electrodes but not for posterior (1.793 vs. 1.679 μV).

To further elaborate the grammaticality X anteriority interaction, separate ANOVAs were conducted for English and Jabberwocky sentences.

English

In English, a grammaticality X anteriority interaction was found [$F(2, 46) = 0.0595$], with ungrammatical sentences less positive than grammatical sentences (1.719 vs. 2.375 μV) over anterior sites but no differences over central (1.856 vs.

2.081 μV) electrodes. This pattern was reversed in the posterior region, with ungrammatical sentences producing greater positivity than grammatical sentences (2.079 vs. 1.674 μV). Although this effect was greater over the left hemisphere, it was not significantly so.

Jabberwocky

A borderline effect of grammaticality was found [$F(1, 23) = 3.177, p = 0.0879$], with grammatical Jabberwocky sentences eliciting greater positivity (2.159 μV) than ungrammatical sentences (1.551 μV). An interaction was found between laterality X anteriority [$F(2, 46) = 8.384, p = 0.0082$], manifesting as significantly more positive amplitudes over the right than the left hemisphere over anterior (2.469 vs. 1.772 μV) and central (2.328 vs. 1.374 μV) regions, but no difference over posterior electrodes (1.537 vs. 1.652 μV).

500-1000 ms (P600)

A main effect of grammaticality was found in this window [$F(1, 23) = 8.939, p = 0.0065$], with ungrammatical sentences producing greater positivity overall than grammatical sentences (2.435 vs. 1.339 μV). A language X anteriority interaction was significant [$F(2, 46) = 6.462, p = 0.0182$], with English sentences resulting in less positive amplitudes than Jabberwocky sentences in anterior electrodes (1.740 vs. 2.173 μV), no differences between languages over central electrodes (2.020 vs. 1.827 μV), and Jabberwocky sentences yielding less positive amplitudes than English sentences over posterior electrodes (1.377 vs. 2.185 μV). A grammaticality X anteriority interaction was also significant [$F(2, 46) = 4.716, p = 0.0404$], driven primarily by the

main effect of grammaticality (see Figure 9). Ungrammatical sentences resulted in increasingly more positive amplitudes than grammatical sentences over anterior (2.214 vs. 1.699 μV), central (2.498 vs. 1.348 μV), and posterior (2.591 vs. 0.971 μV) regions.

To further elaborate the grammaticality X anteriority interaction, separate ANOVAs were conducted for English and Jabberwocky sentences.

English

A main effect of grammaticality was found [$F(1, 23) = 7.554, p = 0.0114$], with ungrammatical English sentences eliciting more positivity than grammatical sentences (2.649 vs. 1.314 μV). A grammaticality X anteriority interaction [$F(2, 46) = 4.868, p = 0.0367$] patterned after the overall effect above (e.g., Figure 9), with ungrammatical English sentences producing increasingly more positive amplitudes than grammatical sentences over anterior (2.002 vs. 1.479 μV), central (2.651 vs. 1.389 μV), and posterior (3.295 vs. 1.075 μV) regions.

Jabberwocky

A main effect of anteriority was seen [$F(2, 46) = 5.924, p = 0.0231$], with amplitudes to Jabberwocky sentences decreasing from anterior (2.173 μV) to central (1.827 μV) to posterior (1.377 μV) regions.

4.4. DISCUSSION

The purpose of this study was to determine whether the (E)LAN and P600, which have been implicated as being sensitive to phrase structure violations, are

equally dependent upon semantic content during auditory sentence processing. The (E)LAN has been suggested to be an early indicator of syntactic processing sensitive to phrase structure violations and independent of semantic content, while the P600 has been described as an index of syntactic (and sometimes semantic) reanalysis or repair following a challenging or violated parse. Comparing ERPs elicited by grammatical vs. ungrammatical sentences in English permits a “baseline” profile of responses for these violations, against which Jabberwocky sentences, which have limited semantic content, can then be compared.

The two phrase structure violations employed in this experiment appear to have been processed differently overall, contrary to what has been reported elsewhere (Canseco-Gonzalez, et al., 1997, in prep). A comparison across Tables 2-4 reveals different areas of significance across the two violation types. In the earliest window (100-300 ms), Jabberwocky sentences of the phrase structure II form elicited amplitudes that were more positive than English sentences; this pattern reversed, however, in the second time window (300-500 ms), with English sentences being more positive than Jabberwocky sentences, but only for phrase structure I sentences. A phrase structure X language X grammaticality interaction in the 100-300 ms window in the omnibus ANOVA revealed no differences in amplitude for grammatical and ungrammatical sentences in English for either phrase structure category. The interaction was instead due to a significant effect of grammaticality on Jabberwocky sentences, with amplitudes to ungrammatical Jabberwocky sentences being considerably more positive than grammatical sentences in phrase structure I

conditions, but the opposite pattern obtaining (i.e., more positive amplitudes to grammatical vs. ungrammatical Jabberwocky sentences) in phrase structure II sentences (see Figure 10).

Such an early effect suggests two possibilities. First, it might be that sentences formed by either phrase structure in English do not provide enough early information for the earliest stages of syntactic processing to be indexed (by an ELAN), while sentences in Jabberwocky do provide enough information for the two phrase structure types to produce distinct results. What seems more likely, however, is that some element of the (acoustic) information presented in Jabberwocky sentences has differential effects before the critical word, producing these differences; perhaps hearing the English verb in a Jabberwocky sentence engages operations which produce a late positivity (P600) that lingers into the analysis epoch for the following word, which is the critical word in this condition. The differences in this earliest window, then, are likely the results of different parsing effects between the two language conditions, making a direct comparison with equivalent English sentences less valid. This possibility is supported by persistent effects of phrase structure observed for the 300-500 and 500-1000 ms windows (see Figure 11), and is perhaps an artifact of the auditory nature of the stimuli in this study.

Because of these differences across phrase structure types, results from each phrase structure type and language will be reported separately below.

Phrase structure I

English

The first effects of grammaticality appear in the 300-500 ms range, with ungrammatical English sentences producing what appears to be an N400 effect. This condition was expected to produce a LAN; although the observed effect is within this time range (and continues somewhat longer), the scalp distribution is not consistent with a typical LAN.

A late positivity elicited by the ungrammatical sentences appears, with a centroparietal maximum, beginning ~800 ms after critical word onset, which then continues throughout the visual epoch. This resembles a traditional (albeit somewhat later) P600, and appears to index attempts to reanalyze or recover structure in this condition compared to grammatical English sentences.

Jabberwocky

An early positivity was seen for ungrammatical Jabberwocky sentences in this window. The functional significance of this is unclear, though it may be a residual positivity “bleeding over” from attempts to parse the English verb one word previously. Notably, no negativities are seen across the epoch. A main effect of grammaticality finds a late positivity emerging at around 800 ms elicited by ungrammatical sentences compared to grammatical sentences. As in the English condition, this appears to be an index of an attempt at recovery or reanalysis, despite the fact that there was a paucity of semantic information available in the Jabberwocky condition. This result was unpredicted by models claiming that the P600 is only elicited in cases in which attempts at meaning are possible, and conflicts somewhat

with results of a previous study using these same materials presented visually (Canseco-Gonzalez, 1997, in prep).

Phrase Structure II

English

A significant result was the finding of an anterior negativity (which was greater over the left hemisphere but not significantly so) to violations in the 300-500 ms window in phrase structure II sentences in English. This strongly resembles a LAN in time course and scalp distribution. Importantly, phrase structure II stimuli were created in such a way that the violation was, in fact, a violation of verb subcategorization rules. Significantly, this result is in line with previous studies employing verb subcategorization violations which elicited a LAN (Osterhout & Holcomb, 1993; Rösler, Friederici, Putz, & Hahne, 1993). This LAN was followed by a broadly distributed positivity to the ungrammatical sentences, having a centroparietal maximum. These results appear to demonstrate that this phrase structure violation also produces a LAN followed by a P600 in English.

Jabberwocky

These results did find an effect of grammaticality (albeit borderline) in the Jabberwocky sentences in this phrase structure type having a distribution consistent with the LAN. Particularly given the results in the English condition, finding a LAN in the Jabberwocky condition appears to suggest that semantic information is not necessarily involved in the process of phrase structure building, even when it involves

violation of verb subcategorization rules. This is consistent with models which posit that assigning thematic roles occurs in a somewhat later stage of parsing LAN (Friederici, 1995, 2002).

A traditional P600 appears in the ungrammatical condition, indicating, as in the English conditions, that semantic information is not necessary for the processes underlying the P600 to be engaged. In other words, the P600 appears to index a process which attempts to reanalyze or reconstitute meaning even when there is a limited amount of semantic information available.

Taken as a whole, these results suggest that structure building and recovery processes during auditory sentence comprehension are not dependent upon the semantic content of the linguistic stream, in contrast to previous results in auditory (Hahne & Jescheniak, 2001) and visual (Canseco-Gonzalez et al., 1997; Münte, Matzke, & Johannes, 1997) presentation modalities.

Interestingly, across phrase structure types, the late positivity effects appear to have a later onset than is typically reported, in both English and Jabberwocky languages; the P600 appears at ~600 ms in PS2 sentences, but it is delayed in PS1 sentences, with an ~800 ms onset. This result may be a symptom of the overall main effect across phrase structure types (beginning with the very early positivity found in PS1 sentences), and may have its source in the different types of critical words being processed across the conditions. That is, the critical word in PS1 sentences indexed a missing element needed to make the previous verb part of an embedded clause; these words were typically nouns or adjectives. The critical word in PS2 sentences was an

English verb that matched (or did not match) the verb subcategorization rules of a preceding English verb. It is therefore possible that the typical P600-indexed repair or reanalysis processes in PS1 sentences (with the missing element) were delayed compared to those in the PS2 sentences. It could be that the parser is sensitive to the need for another verb at this point and there is a pause in the syntax repair process as the incoming word – of an incorrect word class – is “re-checked” for other properties which might make it fit. This more “effortful” parse (compared to the violation in PS2) may result in the delayed P600, though this possibility is only conjecture.

One feature of the late positivities found in the Canseco-Gonzalez et al. (1997) study bears comment. In a visual equivalent to this study, those authors found no P600 effect overall in the Jabberwocky language conditions. The ERPs elicited by the critical words elicited significant positivity in the late time window (500-1000 ms) in both grammatical and ungrammatical Jabberwocky sentences, but because there was no reliable difference between in mean amplitude between these two conditions, the authors concluded that there was no effect of grammaticality as revealed by a P600 in Jabberwocky sentences. Another possible explanation for their result is that the psycholinguistic processes underlying the P600 were operating in both those conditions (e.g., whereas it was not in the English grammatical sentences), leading to no apparent effect across Jabberwocky grammaticality. Given that the current study did find a P600 grammaticality effect in the Jabberwocky sentences, it seems reasonable to invoke an explanation involving task-related processing differences based on features of presentation modality between the two studies.

Finally, the current study attempted to further delineate whether elicitation of the LAN and P600 effects in auditory sentence comprehension depends upon the semantic content of the message. Inconsistent with other authors who have found a LAN followed by a P600 effect in English and Jabberwocky sentences, the current study did not find a LAN for either English or Jabberwocky sentences in either phrase structure violation type. Other authors have found a LAN followed by a P600 elicited by ungrammatical sentences in German and “German Jabberwocky”²⁹ (Hahne & Jescheniak, 2001; Münte, Matzke, & Johannes, 1997). In fact, authors using the identical stimuli (in the visual modality) also found a LAN in both English and Jabberwocky sentences (Canseco-Gonzalez, 1997). The most likely explanation as to why a LAN did not manifest in the present study may be tied to task- or modality-related differences in processing visual and auditory sentences. The present study did, however, find a P600 effect following phrase structure violations in both English and Jabberwocky sentences, a result which, on the surface, suggests that lack of semantic content does not limit the processes underlying the P600. However, because it is difficult to construct “parse-able” materials that are completely devoid of semantic content, there was some semantic information available the stimuli in this study, namely, all the attendant lexicosemantic properties associated with the verbs, adjectives, and function words. Because a certain amount of semantic content was present within the stimuli, it is possible that the parser was engaged in a “ramped-up” version of sentence processing, which resulted in more active processing of the stimuli

²⁹ Jabberwocky sentences produced to conform to word-like characteristics in German. This is significant because German morphosyntax utilizes prefixes and suffixes to carry case marking information which can make phrase structure violations more salient than English equivalents.

in an assertive, “gotta make it work” manner. This would certainly explain the P600 effects seen in Jabberwocky sentences, but it would not fully account for the P600 grammaticality effects seen in both languages whereby ungrammatical sentences required processing that was quantitatively different – as indexed by P600 amplitude – from the processing associated with grammatical sentences. These results therefore do appear to demonstrate that the P600 can be elicited in semantically limited contexts and that attempts to recover meaning are still in effect in sentences with very limited semantic information.

Table 4.1: Mean % correct scores for each of the violation types (Phrase Structure I and Phrase Structure II) across all 24 participants, showing accuracy on two levels of language (English vs. Jabberwocky) and two levels of grammaticality (grammatical vs. ungrammatical). Scores indicate % correct.

	English		Jabberwocky	
	Grammatical	Ungrammatical	Grammatical	Ungrammatical
Phrase Structure 1	80.4	81.6	81.7	82.2
Phrase Structure 2	81.3	82.7	81.7	80.8

Table 4.2: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 100-300 Millisecond Latency Range.

<i>Source</i>	Phrase structure 1		Phrase structure 2	
	<i>F-value</i>	<i>P (HF)</i>	<i>F-value</i>	<i>P (HF)</i>
Language	2.367	.1376	5.202	.0322*
Grammaticality	2.522	.1259	.718	.4054
Laterality	.067	.7983	5.222	.0318*
Anteriority	9.189	.0059**	.634	.5349
Language X Grammaticality	3.563	.0718~	1.842	.1879
Language X Laterality	.335	.5685	.452	.5211
Language X Anteriority	2.550	.1239	1.083	.3472
Grammaticality X Laterality	.507	.4835	.210	.6511
Grammaticality X Anteriority	1.624	.2081	.965	.3887
Laterality X Anteriority	3.770	.0645~	3.640	.0690~
Language X Grammaticality X Laterality	.207	.6534	.050	.8248
Language X Grammaticality X Anteriority	.161	.8520	3.965	.0585~
Language X Laterality X Anteriority	.744	.4808	.766	.4708
Grammaticality X Laterality X Anteriority	.088	.9159	.299	.7430
Language X Grammaticality X Laterality X Anteriority	.325	.7241	.553	.5792

HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1, **p* < 0.05, ***p* < 0.01

Table 4.3: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 300-500 Millisecond Latency Range.

Source	Phrase structure 1		Phrase structure 2	
	F-value	P (HF)	F-value	P (HF)
Language	3.591	.0707~	.118	.7349
Grammaticality	.803	.3794	1.827	.1896
Laterality	.085	.7727	5.370	.0297*
Anteriority	.804	.4536	1.916	.1588
Language X Grammaticality	1.170	.2906	.892	.3546
Language X Laterality	3.552	.0722~	.002	.9680
Language X Anteriority	.663	.5202	.526	.5943
Grammaticality X Laterality	.327	.5730	.296	.5918
Grammaticality X Anteriority	1.751	.1850	3.201	.0868~
Laterality X Anteriority	2.296	.1120	4.563	.0435*
Language X Grammaticality X Laterality	.148	.7041	.280	.6017
Language X Grammaticality X Anteriority	.136	.8735	.974	.3852
Language X Laterality X Anteriority	1.108	.3388	2.191	.1233
Grammaticality X Laterality X Anteriority	.554	.5784	1.255	.2947
Language X Grammaticality X Laterality X Anteriority	2.101	.1339	.133	.8755

HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1, **p* < 0.05, ***p* < 0.01

Table 4.4: Phrase Structure Effects by Language Type: Mean ERP Amplitude ANOVAs in the 500-1000 Millisecond Latency Range.

Source	Phrase structure 1		Phrase structure 2	
	F-value	P (HF)	F-value	P (HF)
Language	.343	.5640	.606	.4444
Grammaticality	.328	.5725	8.939	.0065**
Laterality	1.012	.3249	.532	.4732
Anteriority	.079	.9240	.416	.6624
Language X Grammaticality	.662	.4243	.563	.4606
Language X Laterality	1.024	.3222	.775	.3879
Language X Anteriority	3.804	.0634~	6.462	.0182*
Grammaticality X Laterality	.925	.3462	.813	.3766
Grammaticality X Anteriority	2.182	.1243	4.716	.0404*
Laterality X Anteriority	4.043	.0562~	2.226	.1195
Language X Grammaticality X Laterality	.902	.3521	.022	.8840
Language X Grammaticality X Anteriority	1.153	.3247	1.739	.1870
Language X Laterality X Anteriority	.026	.9748	1.480	.2382
Grammaticality X Laterality X Anteriority	.694	.5046	3.102	.0915~
Language X Grammaticality X Laterality X Anteriority	.109	.8970	.064	.9385

HF = Huynh-Feldt corrected *p* values for *F* tests with more than 1 degree of freedom.

~*p* < 0.1, **p* < 0.05, ***p* < 0.01

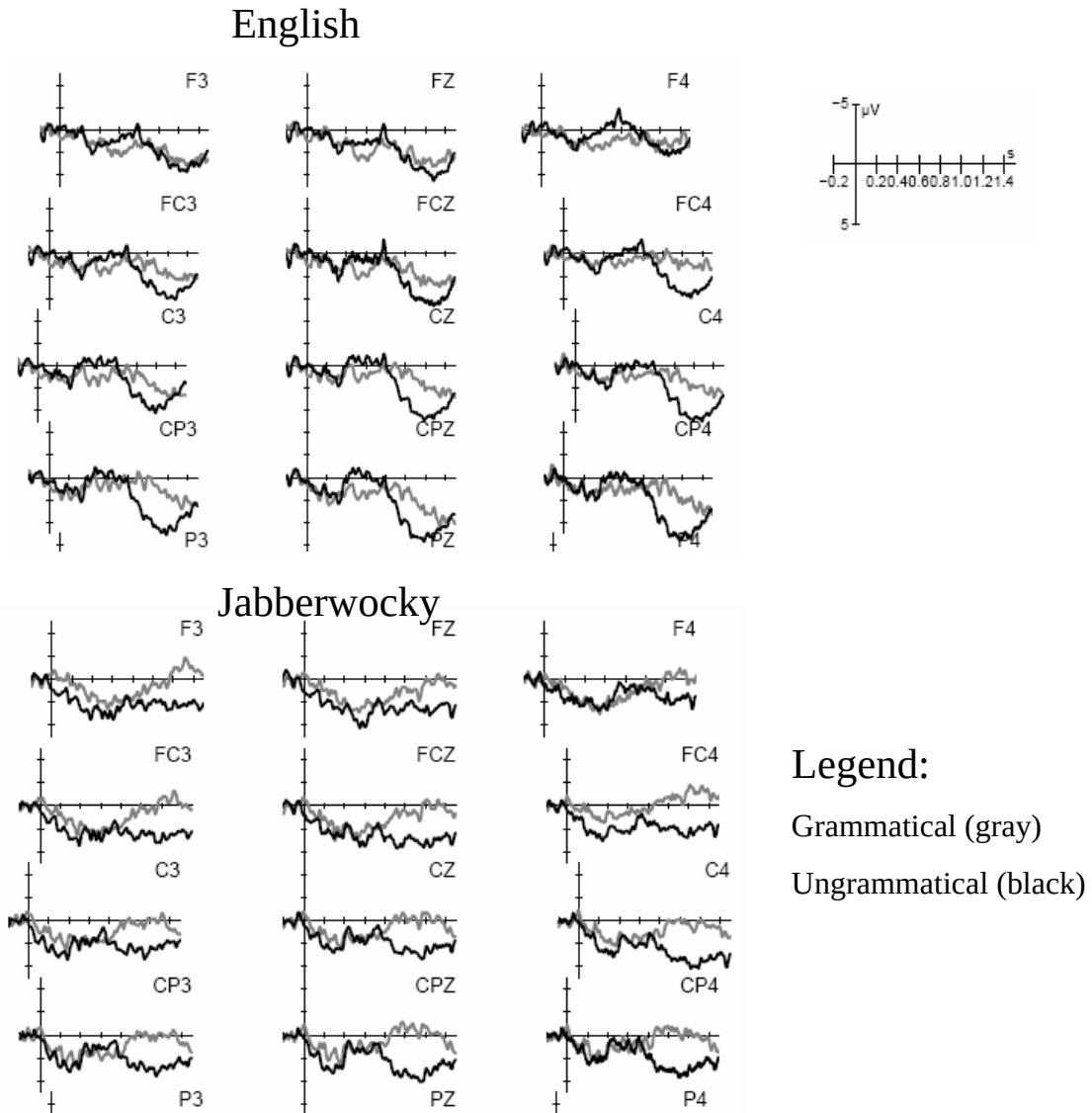


Figure 4.1: Phrase structure I sentences: ERPs elicited by grammatical (gray) and ungrammatical (black) sentences in English (top) vs. Jabberwocky (bottom) are shown over selected electrodes. Amplitudes are in microvolts.

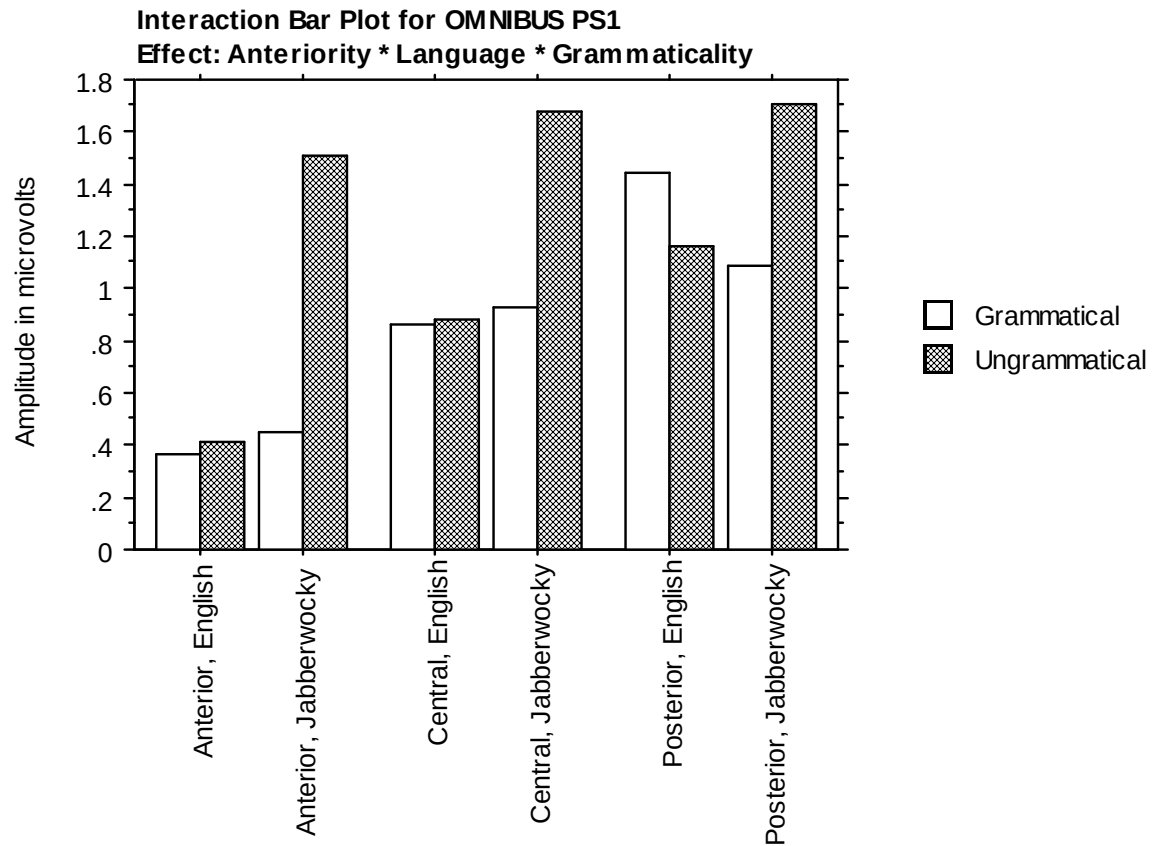


Figure 4.2: Interaction between language and grammaticality in phrase structure I sentences in the 100-300 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (hashed) conditions for English and Jabberwocky sentences are shown; the differences are shown across anterior, central, and posterior electrode groups (for illustrative purposes only).

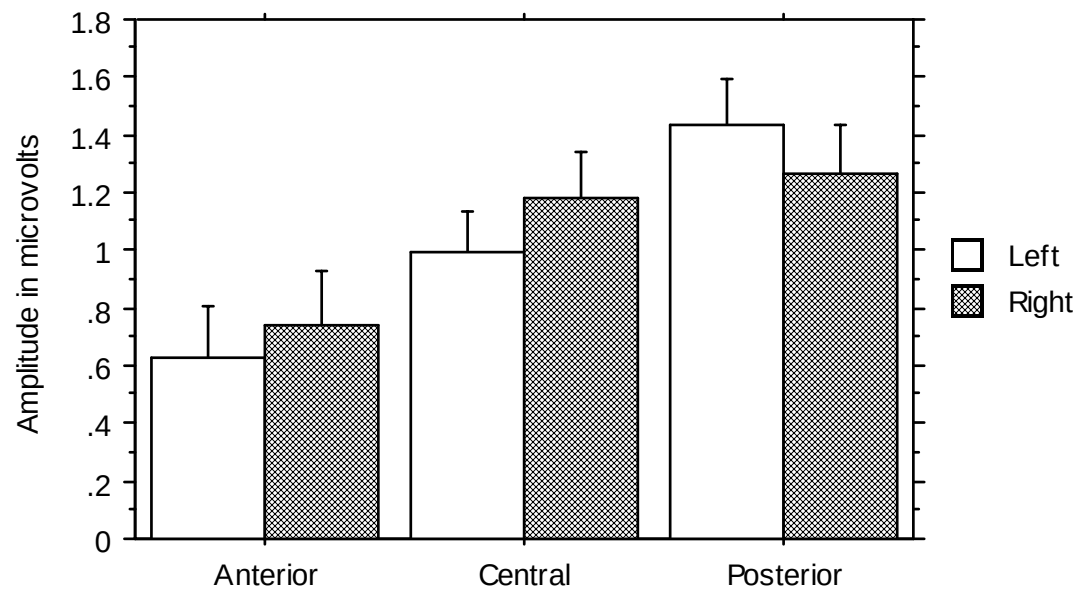


Figure 4.3: Interaction between language X anteriority in phrase structure I sentences in the 100-300 ms window. Mean amplitudes are in microvolts. The difference between left (white) and right (hashed) hemispheres is shown across anterior, central, and posterior electrode groups. Error bars represent standard error.

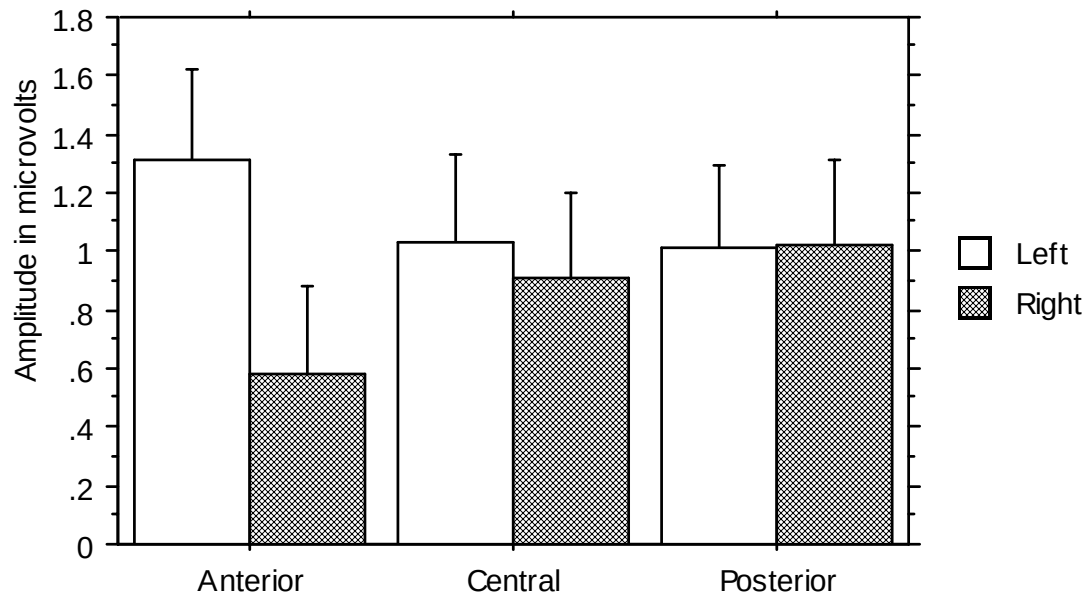


Figure 4.4: Interaction between laterality and anteriority in phrase structure I sentences in the 500-1000 ms window. Mean amplitudes are in microvolts. The difference between left (white) and right (hashed) hemispheres is shown across anterior, central, and posterior electrode groups. Error bars represent standard error.

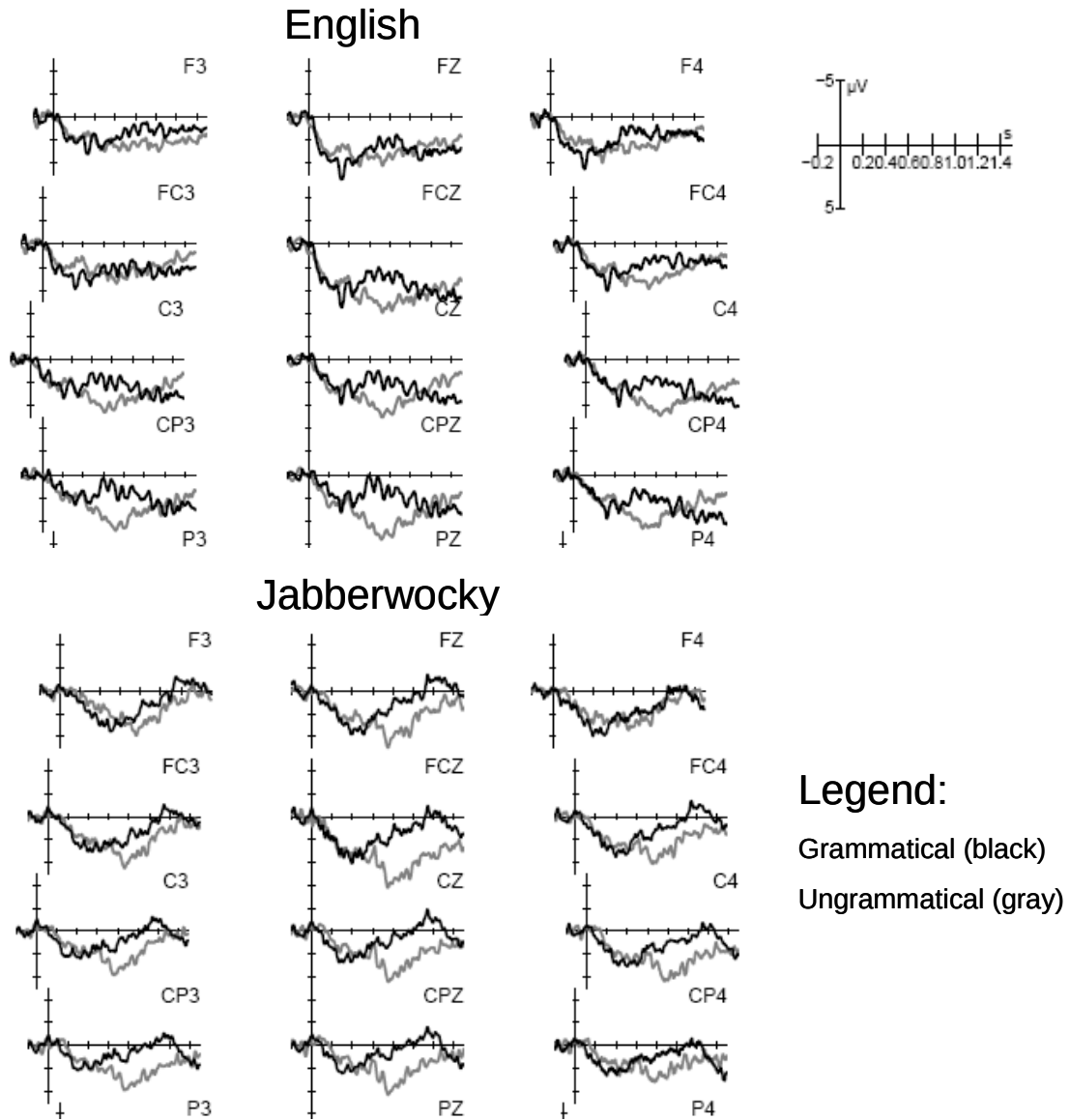


Figure 4.5: Phrase structure II sentences: ERPs elicited by grammatical (black) and ungrammatical (gray) sentences in English (top) vs. Jabberwocky (bottom) are shown over selected electrodes. Amplitudes are in microvolts. (Note that the colors are reversed for grammatical vs. ungrammatical compared to Figure 1.)

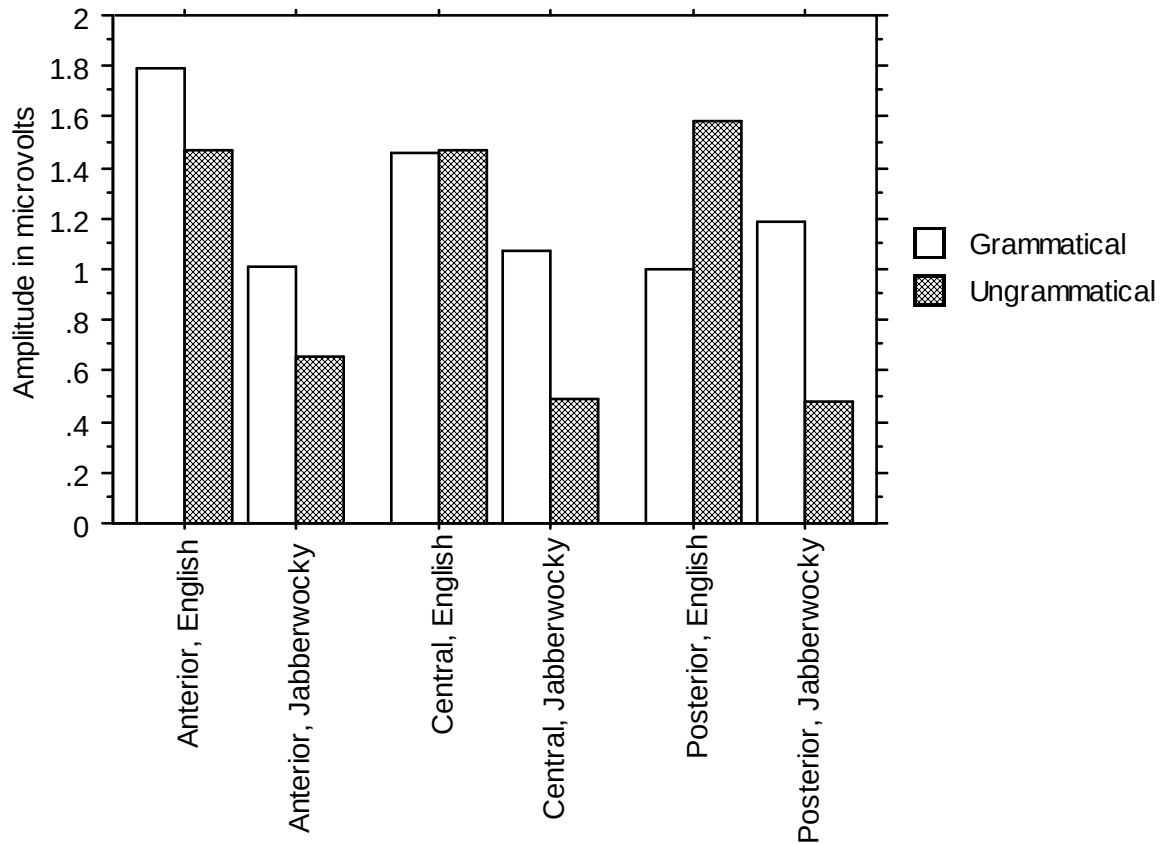


Figure 4.6: Interaction between language X grammaticality X anteriority in phrase structure II sentences in the 100-300 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (crossed) conditions for English and Jabberwocky sentences are shown at anterior, central, and posterior electrode groups.

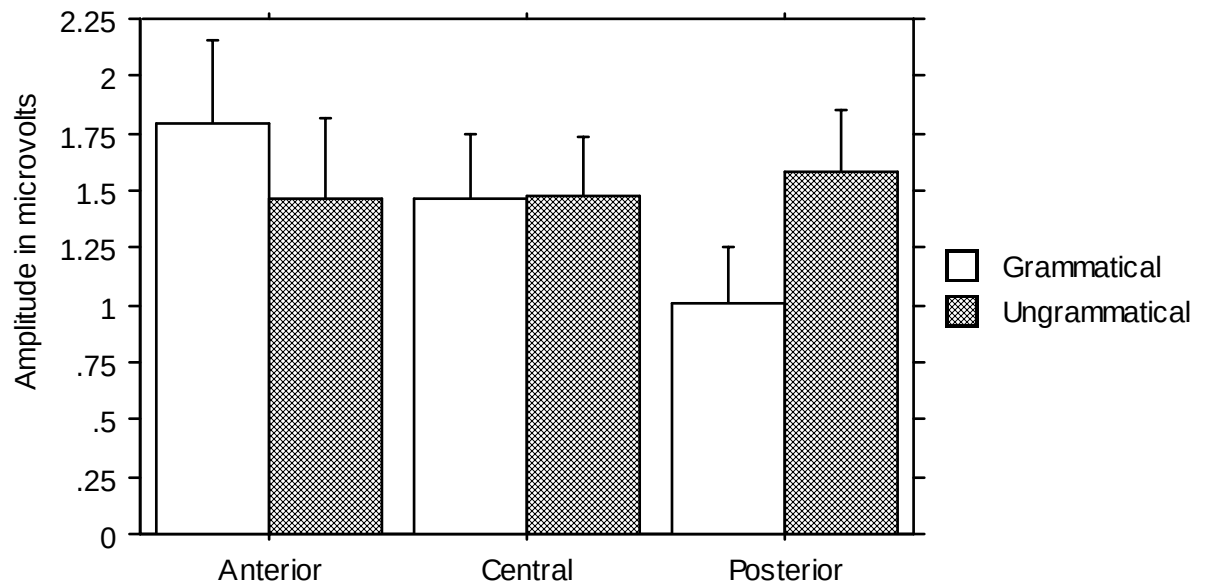


Figure 4.7: Interaction between grammaticality X anteriority in English phrase structure II sentences in the 100-300 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (crossed) conditions are shown at anterior, central, and posterior electrode groups. Error bars represent standard error.

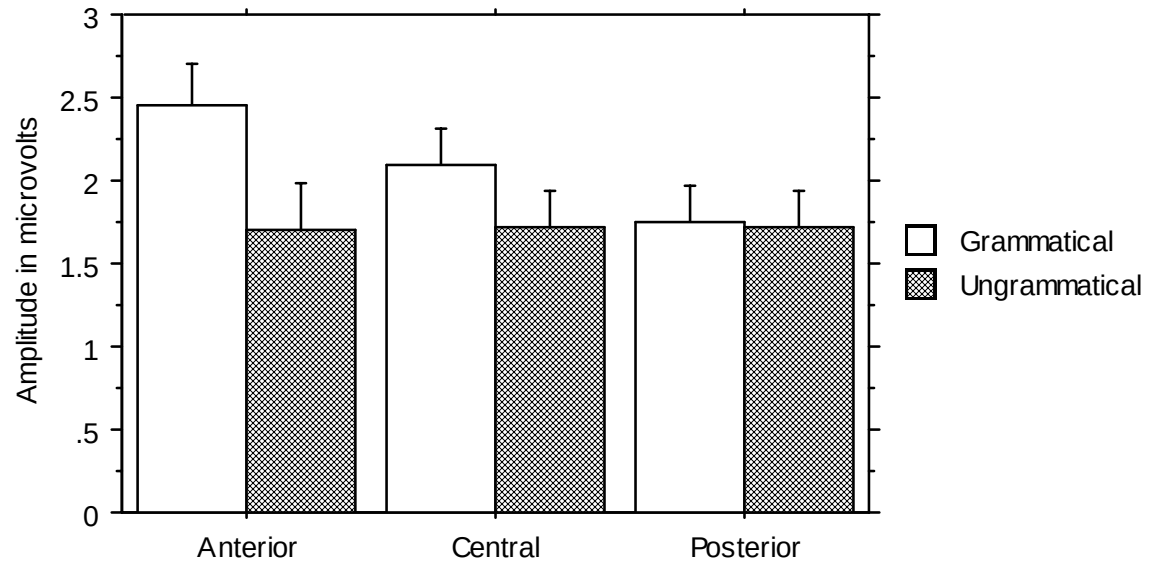


Figure 4.8: Interaction between grammaticality X anteriority in phrase structure II sentences in the 300-500 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (crossed) conditions are shown at anterior, central, and posterior electrode groups. Error bars represent standard error.

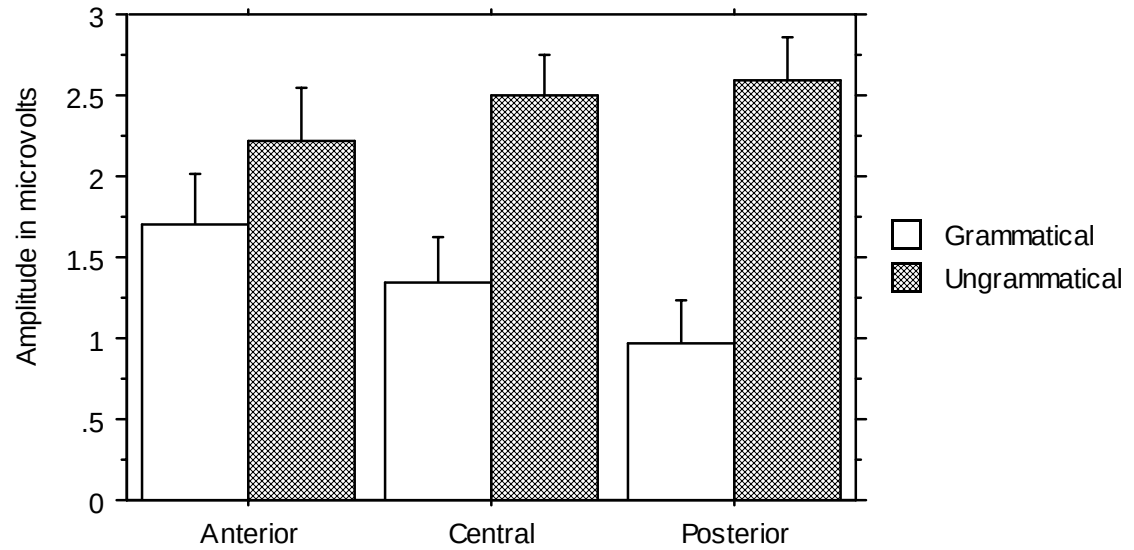


Figure 4.9: Interaction between grammaticality X anteriority in phrase structure II sentences in the 500-1000 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (crossed) conditions are shown at anterior, central, and posterior electrode groups. Error bars represent standard error.

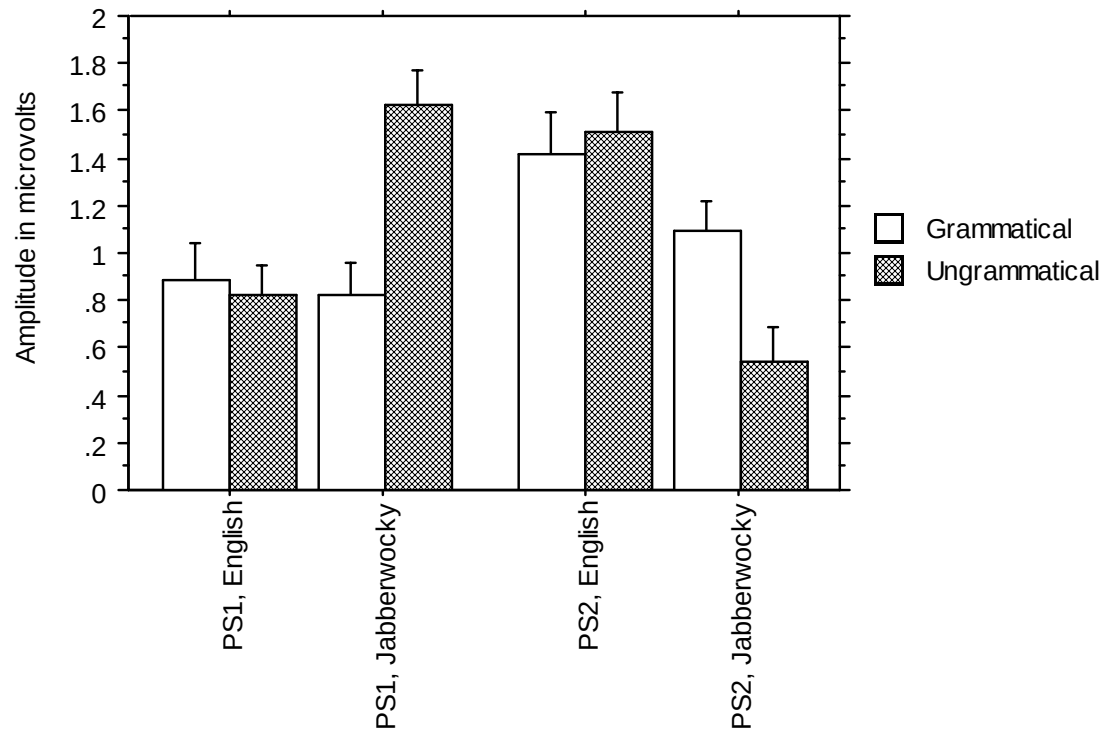


Figure 4.10: Interaction between phrase structure X language X grammaticality in the 100-300 ms window. Mean amplitudes are in microvolts. Grammatical (white) and ungrammatical (crossed) conditions are shown in phrase structure I (PS1) and phrase structure II (PS2) conditions in both English and Jabberwocky languages. Error bars represent standard error.

Phrase Structure I

Phrase Structure II

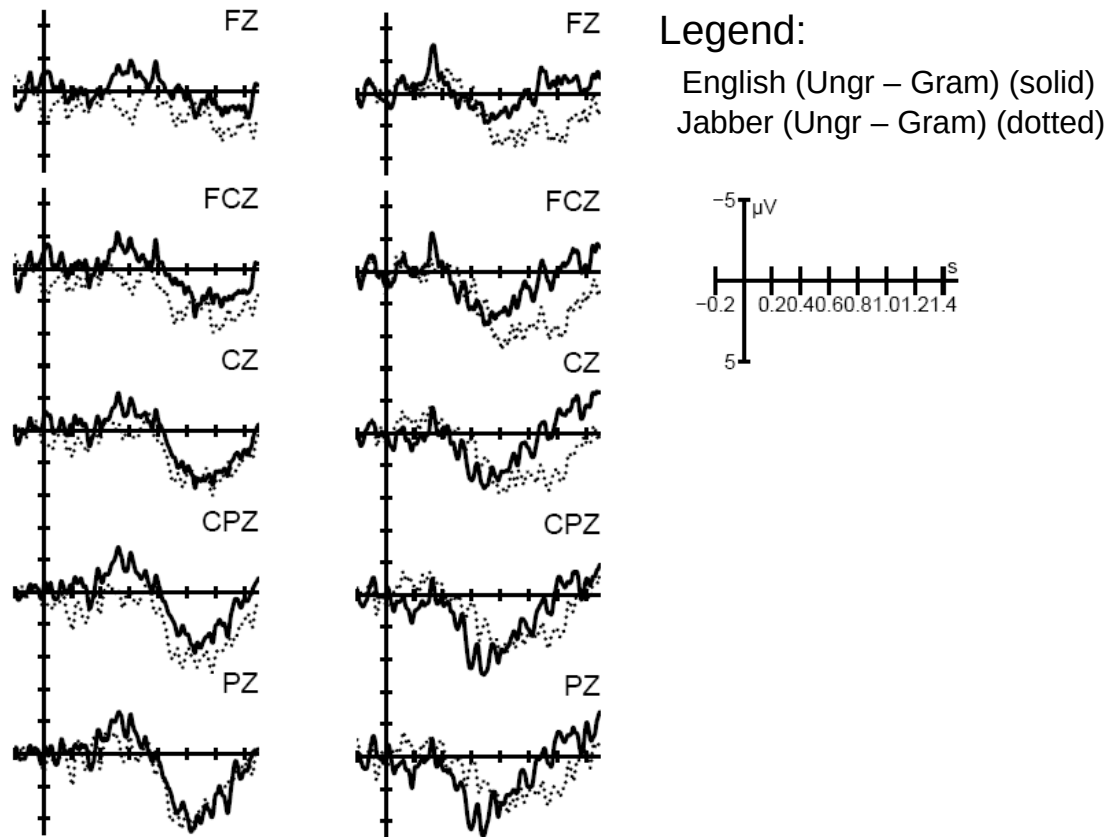


Figure 4.11: Difference ERPs elicited by phrase structure I and phrase structure II sentences shown over selected electrodes. The waves represent subtraction of grammatical waves from ungrammatical in English (solid) and Jabberwocky (dotted) sentences. Amplitudes are in microvolts.

APPENDIX A

Experimental materials for the Jabberwocky study.

PHRASE STRUCTURE I (WORD ORDER)

- 1 a 1 The children that the teacher liked ate pea soup after the show.
- 1 b 1 The children that the teacher liked pea soup after the show.
- 1 c 1 The hessups that the matricle liked ate pea chup after the sholt.
- 1 d 1 The hessups that the matricle liked pea chup after the sholt.
- 2 a 1 The players that the coach hated staged loud riots after the game.
- 2 b 1 The players that the coach hated loud riots after the game.
- 2 c 1 The lufnuts that the easmet hated staged loud sebots after the pob.
- 2 d 1 The lufnuts that the easmet hated loud sebots after the pob.
- 3 a 1 The electricians that the supervisor loathed drank Dutch beer after work.
- 3 b 1 The electricians that the supervisor loathed Dutch beer after work.
- 3 c 1 The killigons that the pessimop loathed drank Dutch gilyn after wezz.
- 3 d 1 The killigons that the pessimop loathed Dutch gilyn after wezz.
- 4 a 1 The professor that the students favored won grant funding for his projects.
- 4 b 1 The professor that the students favored grant funding for his projects.
- 4 c 1 The bellup that the chedons favored won grant kelling for his jepwops.
- 4 d 1 The bellup that the chedons favored grant kelling for his jepwops.
- 5 a 1 The officer that the mafia detested arrested Asian gangs in Chicago.
- 5 b 1 The officer that the mafia detested Asian gangs in Chicago.
- 5 c 1 The veelup that the nasia detested arrested Asian memms in Belio.
- 5 d 1 The veelup that the nasia detested Asian memms in Belio.
- 6 a 1 The rookie that the officers respected captured thirteen criminals in one night.
- 6 b 1 The rookie that the officers respected thirteen criminals in one night.
- 6 c 1 The rebder that the sikclofs respected captured thirteen deminals in one lemm.
- 6 d 1 The rebder that the sikclofs respected thirteen deminals in one lemm.
- 7 a 1 The singer that the girl loved sang sad songs just for her.
- 7 b 1 The singer that the girl loved sad songs just for her.
- 7 c 1 The hizier that the wilk loved sang sad wends just for her.
- 7 d 1 The hizier that the wilk loved sad wends just for her.
- 8 a 1 The soldiers that the general respected earned gold medals of valor.
- 8 b 1 The soldiers that the general respected gold medals of valor.
- 8 c 1 The rensors that the berinal respected earned gold helads of zantor.
- 8 d 1 The rensors that the berinal received gold helads of zantor.
- 9 a 1 The grandmother that the child cherished designed German quilts in her home.
- 9 b 1 The grandmother that the child cherished German quilts in her home.
- 9 c 1 The paritomo that the reg cherished designed German kips in her min.
- 9 d 1 The paritomo that the reg cherished German kips in her min.
- 10 a 1 The clown that the children adored rode huge bicycles in the circus.
- 10 b 1 The clown that the children adored huge bicycles in the cirucs.
- 10 c 1 The bren that the shedron adored rode huge gerters in the nertus.
- 10 d 1 The bren that the shedron adored huge gerters in the nertus.
- 11 a 1 The manicurist that the customer preferred cancelled weekly appointments because of her health.
- 11 b 1 The manicurist that the customer preferred weekly appointments because of her health.
- 11 c 1 The deritatist that the verister preferred cancelled weekly setikastins because of her yill.
- 11 d 1 The deritatist that the verister preferred weekly setikastins because of her yill.
- 12 a 1 The judge that the people endorsed awarded tougher penalties to criminals.
- 12 b 1 The judge that the people endorsed tougher penalties to criminals.

- 12 c 1 The yetts that the nimdan endorsed awarded tougher nerities to serinans.
- 12 d 1 The yetts that the nimdan endorsed tougher nerities to serinans.
- 13 a 1 The candidate that the boss selected borrowed antique furniture for his office.
- 13 b 1 The candidate that the boss selected antique furniture for his office.
- 13 c 1 The deritant that the sholl selected borrowed antique bertwane for his padunt.
- 13 d 1 The deritant that the sholl selected antique bertwane for his padunt.
- 14 a 1 The leader that the students elected chose twelve issues to deal with during the school year.
- 14 b 1 The leader that the students elected twelve issues to deal with during the school year.
- 14 c 1 The selder that the mentobs elected chose twelve hilkers to quilk with during the pim donk.
- 14 d 1 The selder that the mentobs elected twelve hilkers to quilk with during the pim donk.
- 15 a 1 The cousin that the children admired started morning classes in September.
- 15 b 1 The cousin that the children admired morning classes in September.
- 15 c 1 The nimnon that the zellom admired started morning wernses in bigilop.
- 15 d 1 The nimnon that the zellom admired morning wernses in bigilop.
- 16 a 1 The storekeeper that the kids liked bought pink gum in the afternoon.
- 16 b 1 The storekeeper that the kids liked pink gum in the afternoon.
- 16 c 1 The yerterop that the fets liked bought pink vub in the querton.
- 16 d 1 The yerterop that the fets liked pink vub in the querton.
- 17 a 1 The boy that the girl loathed spilled skim milk in the cafeteria.
- 17 b 1 The boy that the girl loathed skim milk in the cafeteria.
- 17 c 1 The sim that the pilk loathed spilled skim nelm in the wertenia.
- 17 d 1 The sim that the pilk loathed skim nelm in the wertenia.
- 18 a 1 The firefighter that the mayor hated fought brush fires at night.
- 18 b 1 The firefighter that the mayor hated brush fires at night.
- 18 c 1 The lelerlit that the chemor hated fought brush zeers at pilt.
- 18 d 1 The lelerlit that the chemor hated brush zeers at pilt.
- 19 a 1 The gardener that the lady favored tended precious flowers at the garden show.
- 19 b 1 The gardener that the lady favored precious flowers at the garden show.
- 19 c 1 The tretitor that the nily favored tended precious klilors at the perto jich.
- 19 d 1 The tretitor that the nily favored precious klilors at the perto jich.
- 20 a 1 The dentist that the woman detested used three forceps to pull out the tooth.
- 20 b 1 The dentist that the woman detested three forceps to pull out the tooth.
- 20 c 1 The tirop that the menow detested used three relons to lell out the numk.
- 20 d 1 The tirop that the menow detested three relons to lell out the numk.
- 21 a 1 The poet that the teacher loved wrote short poems about tragic loss.
- 21 b 1 The poet that the teacher loved short poems about tragic loss.
- 21 c 1 The chium that the verter loved wrote short iluts about ertor nofs.
- 21 d 1 The chium that the verter loved short iluts about ertor nofs.
- 22 a 1 The minister that the villagers preferred allowed Jewish settlers in the church.
- 22 b 1 The minister that the villagers preferred Jewish settlers in the church.
- 22 c 1 The yilister that the burigers preferred allowed Jewish demils in the kerz.
- 22 d 1 The yilister that the burigers preferred Jewish demils in the kerz.
- 23 a 1 The actress that the girl admired created twenty different characters in her career.
- 23 b 1 The actress that the girl admired twenty different characters in her career.
- 23 c 1 The celtron that the dibb admired created twenty alatust beritors in her politom.
- 23 d 1 The celtron that the dibb admired twenty alatust beritors in her politom.
- 24 a 1 The babysitter that the toddler adored refused metal toys from the salesman.
- 24 b 1 The babysitter that the toddler adored metal toys from the salesman.
- 24 c 1 The cheritoster that the delitor adored refused metal yons from the jullmiff.
- 24 d 1 The cheritoster that the delitor adored metal yons from the jullmiff.
- 25 a 1 The preacher that the townspeople cherished lent choir robes to the congregation.
- 25 b 1 The preacher that the townspeople cherished choir robes to the congregation.
- 25 c 1 The chepper that the honspettle cherished lent choir wernts to the yotificals.

- 25 d 1 The chepper that the honspettle cherished choir wernts to the yotificals.
- 26 a 1 The lawyer that the criminal selected gave free consultations once a week.
- 26 b 1 The lawyer that the criminal selected free consultations once a week.
- 26 c 1 The yelom that the virilal selected gave free brentations once a relg.
- 26 d 1 The yelom that the virilal selected free brentations once a relg.
- 27 a 1 The student that the classmates picked proposed secret meetings for the student council.
- 27 b 1 The student that the classmates picked secret meetings for the student council.
- 27 c 1 The nedot that the chertors picked proposed secret shelims for the intant errot.
- 27 d 1 The nedot that the chertors picked secret shelims for the intant errot.
- 28 a 1 The teacher that the principal endorsed praised deaf children for their advancements.
- 28 b 1 The teacher that the principal endorsed deaf children for their advancements.
- 28 c 1 The strenter that the impilac endorsed praised deaf trinints for their desposims.
- 28 d 1 The strenter that the impilac endorsed deaf trinints for their desposims.
- 29 a 1 The candidate that the voters picked fought rough gun restrictions in the Senate.
- 29 b 1 The candidate that the voters picked rough gun restrictions in the Senate.
- 29 c 1 The zadidum that the bemmers picked fought rough ilm kijibons in the wetum.
- 29 d 1 The zadidum that the bemmers picked rough ilm kijibons in the wetum.
- 30 a 1 The president that the people elected named team players to top cabinent posts.
- 30 b 1 The president that the people elected team players to top cabinent posts.
- 30 c 1 The yelitob that the repil elected named team bignums to lup yilito rem.
- 30 d 1 The yelitob that the repil elected team bignums to lup yilito rem.
- 31 a 1 The teacher that the students despised held five students after the bell.
- 31 b 1 The teacher that the students despised five students after the bell.
- 31 c 1 The bigum that the iploofs despised held five pignums after the elp.
- 31 d 1 The bigum that the iploofs despised five pignums after the elp.
- 32 a 1 The astronaut that the admiral appreciated found small asteroids in the solar system.
- 32 b 1 The astronaut that the admiral appreciated small asteroids in the solar system.
- 32 c 1 The gepnot that the axital appreciated found small retiglibs in the welap sinstum.
- 32 d 1 The gepnot that the axital appreciated small retiglibs in the welap sinstum.
- 33 a 1 The gardener that the rancher despised grew vast acres of Indian corn.
- 33 b 1 The gardener that the rancher despised vast acres of Indian corn.
- 33 c 1 The heffiner that the wotter despised grew vast zuknus of hekip suv.
- 33 d 1 The heffiner that the wotter despised vast zuknus of hekip suv.
- 34 a 1 The musician that the woman appreciated played eight songs for her at night.
- 34 b 1 The musician that the woman appreciated eight songs for her at night.
- 34 c 1 The ipnom that the weswut appreciated played eight gers for her at bez.
- 34 d 1 The ipnom that the weswut appreciated eight gers for her at bez.
- 35 a 1 The chauffeur that the lady desired drove slow sportcars at the racetrack.
- 35 b 1 The chauffeur that the lady desired slow sportcars at the racetrack.
- 35 c 1 The fauffer that the ramy desired drove slow pogzars at the yumfift.
- 35 d 1 The fauffer that the ramy desired slow pogzars at the yumfift.
- 36 a 1 The nurse that the man desired raised poor children in her home.
- 36 b 1 The nurse that the man desired poor children in her home.
- 36 c 1 The hess that the mib desired raised poor quepils in her gert.
- 36 d 1 The hess that the mib desired poor quepils in her gert.
- 37 a 1 The child that the woman adopted collected prominent works of major artists.
- 37 b 1 The child that the woman adopted prominent works of major artists.
- 37 c 1 The dex that the jimnum adopted collected prominent perps of himnor jilists.
- 37 d 1 The dex that the jimnum adopted prominent perps of himnor jilists.
- 38 a 1 The kid that the man adopted wanted common things such as a bed.
- 38 b 1 The kid that the man adopted common things such as a bed.
- 38 c 1 The blib that the ven adopted wanted common shoos such as a pij.
- 38 d 1 The blib that the ven adopted common shoos such as a pij.

- 39 a 1 The applicant that counselor rejected threw ice water out the window.
- 39 b 1 The applicant that counselor rejected ice water out the window.
- 39 c 1 The redicant that quizilor rejected threw ice bimils out the jellok.
- 39 d 1 The redicant that quizilor rejected ice bimils out the jellok.
- 40 a 1 The consultant that the president hired noticed angry consumers in the complaints department.
- 40 b 1 The consultant that the president hired angry consumers in the complaints department.
- 40 c 1 The shulatant that the quibiwab hired noticed angry vesvums in the bechooms lelitant.
- 40 d 1 The shulatant that the quibiwab hired angry vesvums in the bechooms lelitant.

PHRASE STRUCTURE II (SUBCATEGORIZATION)

- 41 a 3 The broker persuaded to conceal the transaction was sent to jail.
- 41 b 3 The broker planned to conceal the transaction was sent to jail.
- 41 c 3 The broder persuaded to conceal the tempkishin was sent to julp.
- 41 d 3 The broder planned to conceal the tempkishin was sent to julp.
- 42 a 3 The woman advised to see the play was leaving the theater.
- 42 b 3 The woman agreed to see the play was leaving the theater.
- 42 c 3 The wizlar advised to see the ploob was leaving the zelpof.
- 42 d 3 The wizlar agreed to see the ploob was leaving the zelpof.
- 43 a 3 The journalist encouraged to write the story had missed the deadline.
- 43 b 3 The journalist attempted to write the story had missed the deadline.
- 43 c 3 The figalist encouraged to write the turop had missed the gradvoy.
- 43 d 3 The figalist attempted to write the turop had missed the gradvoy.
- 44 a 3 The grandmother implored to buy the presents had forgotten her purse.
- 44 b 3 The grandmother intended to buy the presents had forgotten her purse.
- 44 c 3 The glofloopen implored to buy the projolp had forgotten her buna.
- 44 d 3 The glofloopen intended to buy the projolp had forgotten her buna.
- 45 a 3 The teacher bribed to steal the money was fined for incompetence.
- 45 b 3 The teacher schemed to steal the money was fined for incompetence.
- 45 c 3 The timshof bribed to steal the mimkoof was fined for humagetence.
- 45 d 3 The timshof schemed to steal the mimkoof was fined for humagetence.
- 46 a 3 The accountant forced to erase the numbers was arrested last week.
- 46 b 3 The accountant declined to erase the numbers was arrested last week.
- 46 c 3 The geshomin forced to erase the tilens was arrested last wulb.
- 46 d 3 The geshomin declined to erase the tilens was arrested last wulb.
- 47 a 3 The athlete permitted to play the game was injured last time.
- 47 b 3 The athlete began to play the game was injured last time.
- 47 c 3 The zumer permitted to play the misk was injured last tuln.
- 47 d 3 The zumer began to play the misk was injured last tuln.
- 48 a 3 The musician allowed to join the orchestra was not very good.
- 48 b 3 The musician aspired to join the orchestra was not very good.
- 48 c 3 The mameshi allowed to join the simitoop was not very blein.
- 48 d 3 The mameshi aspired to join the simitoop was not very blein.
- 49 a 3 The reporter selected to get the story was given a raise.
- 49 b 3 The reporter struggled to get the story was given a raise.
- 49 c 3 The tiranen selected to get the turop was given a yeld.
- 49 d 3 The tiranen struggled to get the turop was given a yeld.
- 50 a 3 The swimmer urged to lose weight was beginning a diet.
- 50 b 3 The swimmer decided to lose weight was beginning a diet.
- 50 c 3 The klupter urged to lose blim was beginning a dinug.
- 50 d 3 The klupter decided to lose blim was beginning a dinug.
- 51 a 3 The mechanic trusted to repair the car had quit his job.
- 51 b 3 The mechanic refused to repair the car had quit his job.
- 51 c 3 The oremy trusted to repair the voon had quit his flib.

- 51 d 3 The oremy refused to repair the voon had quit his flib.
- 52 a 3 The dentist invited to meet the actress was nervous last night.
- 52 b 3 The dentist hoped to meet the actress was nervous last night.
- 52 c 3 The undzer invited to meet the shulog was nervous last zuld.
- 52 d 3 The undzer hoped to meet the shulog was nervous last zuld.
- 53 a 3 The nurse induced to leave the patient was reprimanded very severely.
- 53 b 3 The nurse hesitated to leave the patient was reprimanded very severely.
- 53 c 3 The kolm induced to leave the gevorn was reprimanded very sunikly.
- 53 d 3 The kolm hesitated to leave the gevorn was reprimanded very sunikly.
- 54 a 3 The policeman ordered to arrest the man had hurt his hand.
- 54 b 3 The policeman started to arrest the man had hurt his hand.
- 54 c 3 The bobean ordered to arrest the hoont had hurt his ferg.
- 54 d 3 The bobean started to arrest the hoont had hurt his ferg.
- 55 a 3 The electrician hired to install the light had a good reputation.
- 55 b 3 The electrician tried to install the light had a good reputation.
- 55 c 3 The fufstishen hired to install the tez had a good agaploshen.
- 55 d 3 The fufstishen tried to install the tez had a good agaploshen.
- 56 a 3 The executive ordered to balance the budget was fired for incompetence.
- 56 b 3 The executive planned to balance the budget was fired for incompetence.
- 56 c 3 The shoytemel ordered to balance the boyem was fired for humagetence.
- 56 d 3 The shoytemel planned to balance the boyem was fired for humagetence.
- 57 a 3 The politician urged to run for office was meeting with voters.
- 57 b 3 The politician decided to run for office was meeting with voters.
- 57 c 3 The zashenin in urged to run for vinter was meeting with entrobs.
- 57 d 3 The zashenin in decided to run for vinter was meeting with entrobs.
- 58 a 3 The student allowed to have a party had bought new albums.
- 58 b 3 The student intended to have a party had bought new albums.
- 58 c 3 The shalik allowed to have a velder had bought new hunzels.
- 58 d 3 The shalik intended to have a velder had bought new hunzels.
- 59 a 3 The spy advised to tell the truth had lied for years.
- 59 b 3 The spy started to tell the truth had lied for years.
- 59 c 3 The nup advised to tell the trid had lied for nors.
- 59 d 3 The nup started to tell the trid had lied for nors.
- 60 a 3 The repairman trusted to fix the television was selling bad parts.
- 60 b 3 The repairman began to fix the television was selling bad parts.
- 60 c 3 The gorniklom trusted to fix the visiposin was selling bad klups.
- 60 d 3 The gorniklom began to fix the visiposin was selling bad klups.
- 61 a 3 The senator bribed to sell the secrets was arrested for espionage.
- 61 b 3 The senator schemed to sell the secrets was arrested for espionage.
- 61 c 3 The kelmuzer bribed to sell the maziks was arrested for qualiafage.
- 61 d 3 The kelmuzer schemed to sell the maziks was arrested for qualiafage.
- 62 a 3 The scientist selected to win the prize had arrived by plane.
- 62 b 3 The scientist aspired to win the prize had arrived by plane.
- 62 c 3 The beyamer selected to win the murm * arrived by klont.
- 62 d 3 The beyamer aspired to win the murm * arrived by klont.
- 63 a 3 The referee permitted to make the decision had to think fast.
- 63 b 3 The referee tried to make the decision had to think fast.
- 63 c 3 The flamiter permitted to make the lustishken had to think mest.
- 63 d 3 The flamiter tried to make the lustishken had to think mest.
- 64 a 3 The electrician hired to repair the furnace had finished the job.
- 64 b 3 The electrician attempted to repair the furnace had finished the job.
- 64 c 3 The fuferstishen hired to repair the safirn had finished the jip.
- 64 d 3 The fuferstishen attempted to repair the safirn had finished the jip.

- 65 a 3 The shopper encouraged to buy the coat was given a discount.
- 65 b 3 The shopper declined to buy the coat was given a discount.
- 65 c 3 The gezung encouraged to buy the kib was given a tatel.
- 65 d 3 The gezung declined to buy the kib was given a tatel.
- 66 a 3 The chairman persuaded to answer the question was preparing to resign.
- 66 b 3 The chairman refused to answer the question was preparing to resign.
- 66 c 3 The shofen persuaded to answer the filsoof was preparing to murtig.
- 66 d 3 The shofen refused to answer the filsoof was preparing to murtig.
- 67 a 3 The athlete induced to sign the contract had injured his leg.
- 67 b 3 The athlete hoped to sign the contract had injured his leg.
- 67 c 3 The zumer induced to sign the fargan had injured his loz.
- 67 d 3 The zumer hoped to sign the fargan had injured his loz.
- 68 a 3 The actress implored to learn her lines was ready to quit.
- 68 b 3 The actress struggled to learn her lines was ready to quit.
- 68 c 3 The madyl implored to learn her woffs was ready to nept.
- 68 d 3 The madyl struggled to learn her woffs was ready to nept.
- 69 a 3 The student forced to do the assignment was failing the course.
- 69 b 3 The student hesitated to do the assignment was failing the course.
- 69 c 3 The shalik forced to do the frayleker was failing the blep.
- 69 d 3 The shalik hesitated to do the frayleker was failing the blep.
- 70 a 3 The librarian invited to give a speech had drunk too much.
- 70 b 3 The librarian agreed to give a speech had drunk too much.
- 70 c 3 The ferizian invited to give a smool had drunk too much.
- 70 d 3 The ferizian agreed to give a smool had drunk too much.
- 71 a 3 The janitor pressured to fix the appliance was working all day.
- 71 b 3 The janitor sought to fix the appliance was working all day.
- 71 c 3 The kilburn pressured to fix the pligalance was working all mib.
- 71 d 3 The kilburn sought to fix the pligalance was working all mib.
- 72 a 3 The man obliged to rebuild the house was told to stop.
- 72 b 3 The man wished to rebuild the house was told to stop.
- 72 c 3 The vilb obliged to rebuild the glent was told to blep.
- 72 d 3 The vilb wished to rebuild the glent was told to blep.
- 73 a 3 The researcher compelled to tell the lie was denied the funding.
- 73 b 3 The researcher desired to tell the lie was denied the funding.
- 73 c 3 The jupiltop compelled to tell the hig was denied the zurping.
- 73 d 3 The jupiltop desired to tell the hig was denied the zurping.
- 74 a 3 The bookie coerced to fix the bet was caught by the police.
- 74 b 3 The bookie yearned to fix the bet was caught by the police.
- 74 c 3 The silpy coerced to fix the vit was caught by the injut.
- 74 d 3 The silpy yearned to fix the vit was caught by the injut.
- 75 a 3 The surgeon entrusted to perform the surgery had left town yesterday.
- 75 b 3 The surgeon offered to perform the surgery had left town yesterday.
- 75 c 3 The lijug entrusted to perform the getipy had left town yesterday.
- 75 d 3 The lijug offered to perform the getipy had left town yesterday.
- 76 a 3 The lawyer told to win the trial had lost the case.
- 76 b 3 The lawyer ventured to win the trial had lost the case.
- 76 c 3 The kimpler told to win the zirtel had lost the fid.
- 76 d 3 The kimpler ventured to win the zirtel had lost the fid.
- 77 a 3 The jockey commanded to ride the horse had fallen down suddenly.
- 77 b 3 The jockey aimed to ride the horse had fallen down suddenly.
- 77 c 3 The mikley commanded to ride the nerp had fallen down blefenly.
- 77 d 3 The mikley aimed to ride the nerp had down blefenly.
- 78 a 3 The librarian forced to reshelve the books had left the library.

- 78 b 3 The librarian promised to reshelve the books had left the library.
- 78 c 3 The migalagop forced to reshelve the bifs had left the digalry.
- 78 d 3 The migalagop promised to reshelve the bifs had left the digalry.
- 79 a 3 The programmer authorized to write the manual had finished under budget.
- 79 b 3 The programmer proposed to write the manual had finished under budget.
- 79 c 3 The culnipert authorized to write the sinitif had finished under kulnid.
- 79 d 3 The culnipert proposed to write the sinitif had finished under kulnid.
- 80 a 3 The author driven to finish the novel was promised 5,000 dollars.
- 80 b 3 The author anticipated to finish the novel was promised 5,000 dollars.
- 80 c 3 The quilmpop driven to finish the higlup was promised 5,000 aglits.
- 80 d 3 The quilmpop anticipated to finish the higlup was promised 5,000 aglits.

REFERENCES

- Ainsworth-Darnell, K, Shulman, HG, & Boland, JE (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. *Journal of Memory and Language*, 38, 112-130.
- Balconi, M, & Pozzoli, U (2005). Comprehending semantic and grammatical violations in Italian: N400 and P600 comparison with visual and auditory stimuli. *Journal of Psycholinguistic Research*, 34, 71-98.
- Canseco-Gonzalez, E, Swinney, D, Love, T, Walenski, M, Ahrens, K, & Neville, H (1997). Processing grammatical information using Jabberwocky sentences: an ERP study. Fourth Annual Meeting of the Cognitive Neuroscience Society Boston, MA.
- De Vincenzi, M, Job, R, Di Matteo, R, Angrilli, A, Penolazzi, B, Ciccarelli, L, & Vespignani, F (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain & Language*, 85, 280-296.
- Featherston, S, Gross, M, Münte, TF, & Clahsen, H (2000). Brain potentials in the processing of complex sentences: an ERP study of control and raising constructions. *Journal of Psycholinguistic Research*, 29, 141-154.
- Fiebach, CJ, Schlesewsky, M, & Friederici, AD (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, 47, 250-272.
- Friederici, AD, Mecklinger, A, & Hahne, A (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1219-1248.
- Gunter, TC, Friederici, AD, & Schriefers (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556-568.
- Gunter, TC, Stowe, LA, & Mulder, G (1997). When syntax meets semantics. *Psychophysiology*, 34, 660-676.
- Hahne, A & Friederici, AD (1998). ERP evidence for autonomous first-pass parsing processes in auditory language comprehension. *Journal of Cognitive Neuroscience (Suppl.)*, S125.

- Hahne, A & Friederici, AD (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11, 194-205.
- Hahne & Friederici (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13, 339-356.
- Kaan, E, Harris, A, Gibson, E, & Holcomb, P (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201.
- MacDonald, M, Pearlmutter, N, & Seidenberg, M (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.
- Münte, TF, Matzke, M, & Johannes, S (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience*, 9, 319-329.
- Neville, HJ, Nicol, JL, Barss, A, Forster, KI, & Garrett, MF (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3, 151-165.
- Osterhout, L & Nicol, J (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283-317.
- Phillips, C, Kazanina, N, Wong, K, & Ellis, R (2001). ERP evidence on the time course of processing demands in wh-dependencies. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing, Philadelphia.
- van Herten, M, Kolk, HH J, & Chwilla, DJ (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22, 241–255.

Chapter 5

Conclusion

This dissertation presented a series of experiments examining the availability and time course of processing different lexical properties during on-going auditory sentence comprehension. These experiments exploited lexical properties on a continuum – from syntactically relevant properties (phrase structure violations without semantic context), to properties of verbs whose interpretation hinges on the semantics of subsequent sentential information (aspectual coercion), to properties that rely primarily on a semantic representation (concreteness). These experiments represent a survey of a range of properties and the time at which these properties are available during real time auditory sentence processing.

The first of these experiments used the lexical property of “concreteness” – a lexical-conceptual notion – in nouns and their coreferent pronouns. This experiment was designed to test whether, and at what point in time, such semantic lexical properties as concreteness are available during the coreference process. The second experiment investigated the lexical property of verb aspect under the influence of aspectual coercion, in which a mismatch exists between the temporal markedness of a verb (single time vs. continuous) and a complement modifier. The purpose of this experiment was to determine whether the process of aspectual coercion can be indexed neurophysiologically as being a semantic process, and whether that process takes place immediately at the point when coercion is licensed. The third experiment examined the reliance of structure-building (syntactic) verb properties on semantic

content. This experiment was designed to examine whether violations of phrase structure (expected word order and verb subcategorization preferences) generate the same (early and late) syntactic reflexes in English sentences as they do in sentences with impoverished lexicosemantic content (i.e., “Jabberwocky” sentences).

The results of the first experiment indicate that an antecedent’s semantic lexical property of concreteness is available immediately at a coreferent pronoun. Because concreteness is viewed as indexing conceptual properties of a word’s representation, these results are taken as support for models in which antecedent reactivation is accomplished at a “deep”, conceptual level. The early availability of such a semantic property of an antecedent suggests that, although perhaps not necessary for a successful parse, semantic lexical information of an antecedent may be consulted during the coreference process.

Results from the second experiment suggest that properties involved in the aspectual coercion process are engaged immediately upon encountering a coercion-licensing modifier, rather than being postponed to some later time. These data also indicate that the neurophysiological correlates of coercion processes may be similar in point action verbs and process verbs, a result that is inconsistent with claims that increased processing costs associated with coercion are due to reinterpretation and generation of an iterative, repetitive representation. Significantly, no ERPs signatures consistent with structural reprocessing (P600) were found in the enrichment conditions, suggesting that this process, indeed, may be primarily semantic.

The third experiment found that both phrase structure violations in both English and Jabberwocky sentences produced a P600 effect, an effect typically associated with structural reanalysis or attempts at meaning recovery. No evidence of early phrase structure building (predicted to be indexed by a LAN) was found. Given that violations in Jabberwocky sentences produced a P600 – as in English violations – it appears that semantic lexical properties of a verb’s arguments are not necessary to engage the cognitive mechanisms underlying the P600. Compared to previous investigations which found a LAN elicited by both types of phrase structure violations in English, German, and Jabberwocky sentences, failure to find a LAN in this study suggests that there may be a difference in the way comprehenders process language between visual and auditory modalities. The finding of time course differences across violation types, however, underscores the need for careful stimulus selection in psycholinguistic research.

Taken together, these studies provide important insights into the temporal and psycholinguistic nature of on-line sentence processing. Overall, these results support the relative independence of lexical from structural properties, in opposition to lexically-based constraint-satisfaction accounts. However, the fact that semantically-relevant properties are available quite early during sentence processing suggests that accounts prohibiting the influence of semantic properties in the early stages of processing may be excessively limiting. That is, in structurally unambiguous constructions of the sort investigated in the experiments described in this dissertation, semantic lexical properties do appear to be used early, and perhaps even initially,

during real-time auditory sentence processing routines. The study that was predicted to show that structural properties are engaged in the absence of semantic properties failed to provide evidence for early structural processing, although it did find what may be attempts at structural reprocessing in stimuli with limited semantic content. These patterns suggest a potential architecture for a sentence processing system that specifies a *lexically-driven, but structurally independent* system in which certain types of lexical information are made available at principled (and temporally distributed) points during the unfolding of the sentence.

Use of the ERP methodology in these studies permits a fine-grained examination of the time course of processing of these properties; these results indicate (particularly in the first and second experiments) that the brain is sensitive to lexical properties well before previous reaction time data have been able to demonstrate (i.e., prior to post-processing effects associated with secondary tasks). Concerned with different lexical properties engaged in distinct psycholinguistic processes, these studies provide evidence in a variety of contexts on the immediate nature of on-line lexical processing.