Title
Electrophysiological explorations of linguistic pre- activation and its consequences during online sentence processing

Author
DeLong, Katherine Ann

Publication Date
2009

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Electrophysiological explorations of linguistic pre-activation and its consequences during online sentence processing

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

in

Cognitive Science

by

Katherine Ann DeLong

Committee in charge:
Professor Marta Kutas, Co-Chair
Professor Jeffrey Elman , Co-Chair
Professor Seana Coulson
Professor Victor Ferreira
Professor Steven Hillyard

2009

The dissertation of Katherine Ann DeLong is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California, San Diego

2009

DEDICATION

Dedicated to

Dimitri Tsintikidis

whose joy in his work made a
scientific career seem appealing

EPIGRAPH

Chance favors only those minds which are prepared.

Louis Pasteur

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

Chapter 4

Chapter 5

ACKNOWLEDGEMENTS

It is hard to imagine this dissertation could have been written if not for one person – my advisor, Marta Kutas. She graciously accepted me into her lab, though I was green beyond belief, both in areas of language processing and scientific methodology. Her encouraging words, that the field of Cognitive Science takes all kinds, gave me hope that a transition into experimental work was feasible. Her generous offerings of time, support (personal and financial), opportunities and mentoring extended well beyond what most graduate students receive from their advisors. Her constraints in choosing areas for investigation were few, but always with the proviso to do thoughtful and thorough science. I could not have hoped for a better place to have landed. Thank you, Marta, for everything.

I would also like to thank Jeff Elman for agreeing to co-chair my dissertation committee – amidst his numerous roles and responsibilities – as well as making time for me whenever I have needed it. I am also grateful to Seana Coulson, particularly for the Foundations course she taught during my first year. That class introduced me to many topics that I continue to puzzle and speculate about today. Seana's availability and advice throughout the years have additionally been extremely valuable. Many thanks, too, to Vic Ferreira and Steve Hillyard for serving on my committee and providing their feedback and expertise.

I am also obliged to all past and present members of the Kutas Lab. In particular, I would like to thank Tom Urbach for his training, collaboration and for being a great sounding board. I am extremely grateful to Esmeralda de Ochoa, whose rigor and professionalism in data collection and all things lab-related I can only aspire to, and whose steadfast friendship has been invaluable. Thanks, too, to David Groppe for being such a great lab mate, and for all the help with my statistical questions. And to Nicole Wicha, Kara Federmeier and Pia Knoeferle, your girl power has inspired me! My thanks extend to

Michael Kiang, too, for his scientific companionship, especially during those years when it was just the two of us working in the back room! And last but not least, I am indebted to Paul Krewski, who resuscitated me from computational disasters on countless occasions.

I would also like to express my gratitude to CRL and INC for both their financial support over the years as well as allowing me to be part of their rich communities of cognitive scientists. Many thanks, too, to Beverley Walton for coordinating the business end of things so seamlessly and graciously.

Portions of Chapter 1 of this thesis (descriptions of the ERP methodology) were extracted from a chapter co-written by myself along with Marta Kutas and Michael Kiang, to be included in the *Handbook of Psycholinguistic & Cognitive Processes: Perspectives in Communication Disorders* (Eds. Jackie Guendouzi, Filip Loncke & Mandy J. Williams, Psychology Press, in press). A significant portion of Chapter 2 is comprised of material from an article in *Nature Neuroscience*, Volume 8(8), Aug 2005, 1117-1121, DeLong, Katherine A., Urbach, Thomas P. & Kutas, Marta, for which I was the primary author and investigator.

Finally, I could not have made it through this journey without the support of my friends and family. A special thanks to Janet Sprintall and Alexander (Sasha) Gershunov for helping open the door to an academic life in the first place. I am infinitely grateful to my sister, Beth DeLong, for her constant companionship, sunny disposition, easy laughter and ready ear-lending. And to my two children, my wonderful life-changers, Thaleia and Xander, who have elevated the goal of working efficiently, so that I can get home in good time to see their shining faces. Lastly, whatever I write here cannot even come close to expressing how much the encouragement and support of my husband, Dimitri Tsintikidis, have meant to me. His positive attitude, steady partnership, and belief in (and tolerance of) me through the ups and downs have not only sustained me, but made this whole endeavor meaningful.

VITA

1992	Bachelor of Music with High Distinction, minor in German University of Iowa
2005	Master of Science in Cognitive Science University of California, San Diego (UCSD)
2009	Doctor of Philosophy in Cognitive Science University of California, San Diego (UCSD)

Publications

DeLong, K.A., Urbach, T.P. & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117-21.

Kutas, M. & DeLong, K.A. (2008). A Sampler of Event-Related Brain Potential (ERP) Analyses of Language Processing. In Zvia Breznitz (Ed.), *Brain Research in Language*. (pp. 153-186). New York: Springer-Verlag.

Kutas, M., DeLong, K.A. & Kiang, M. (In press). From phonemes to discourse: Event-related brain potential (ERP) componentry and paradigms for investigating normal and abnormal language processing. In Jacqueline Guendouzi (Ed.). *Handbook of Psycholinguistics and Cognitive Processing: Perspectives in Communication Disorders*. Psychology Press.

Abstracts and presentations

DeLong, K.A. & Kutas, M. Brainwave studies of contextual preactivation of lexical word forms. Symposium presentation at Society for Psychophysiological Research (SPR) 48th Annual Meeting. October 1-5, 2008, Austin, TX.

DeLong, K.A., Urbach, T.P. & Kutas, M. (2007). A cost to mispredicting: Effects of sentential constraint violations. Poster presentation at 20th Annual CUNY Conference on Human Sentence Processing, March 29-31, 2007, La Jolla, CA.

DeLong, K.A. & Kutas, M. A graded ERP frontal positivity indexing expectation for a part of speech. Poster presentation at Society for Psychophysiological Research (SPR) 46th Annual Meeting, October 25 - 29, 2006, Vancouver, BC, Canada.

DeLong, K.A. (2005). ERP investigations of word pre-activation during sentence comprehension. University of Maastricht, Department of Neurocognition Colloquium, Maastricht, The Netherlands, December 9, 2005.

DeLong, K.A., Urbach, T.P. & Kutas, M. (2005). Electrophysiological evidence for pre-activation of specific word forms during language comprehension. Poster presentation at 12th Annual Meeting of the Cognitive Neuroscience Society, April 10 - 12, 2005, New York, NY.

DeLong, K.A., Urbach, T.P. & Kutas, M. (2005). Pre-activation of specific word forms during language comprehension. Poster presentation at Helsinki Winter School in Cognitive Neuroscience, March 2-11, 2005, University of Helsinki, Finland.

ABSTRACT OF THE DISSERTATION

Electrophysiological explorations of linguistic pre-activation
and its consequences during online sentence processing

by

Katherine Ann DeLong

Doctor of Philosophy

University of California, San Diego, 2009

Professors Marta Kutas and Jeffrey Elman, Co-Chairs

Historically, anticipation has played only a minor role in language comprehension theories, with the potential for error precluding the utility of such a strategy, except in conditions of high constraint or degraded input. Such views, however, may not factor in how meaning is stored, retrieved or constructed in the brain. Recently, evidence for the benefits (and to a lesser extent, the costs) of linguistic prediction has begun to accrue, with methodologically diverse findings revealing that various aspects of comprehension (e.g., semantic, syntactic, and phonological) are shaped by top-down processing.

Two general difficulties of investigating predictive sentence comprehension are that (a) effects of anticipation and integration are difficult to disentangle, and (b) finding evidence for upcoming events requires clever manipulations and experimental techniques

that allow for continuous tracking of responses as contexts unfold. The research presented in this dissertation addresses these concerns by using paradigms (and methods) that allow for pre-target detection of prediction as well as monitoring throughout subsequently predicted targets. We utilize event-related brain potentials (ERPs) and a variety of analytical approaches to investigate probabilistic preactivation across a range of constraining sentence contexts. These studies contrast the contributions of facilitative context (a narrowing of constraint) with semantic fit of particular words. They also examine the consequences of unrealized linguistic predictions – i.e., contextually constrained information presumably preactivated, but ultimately disconfirmed by the input.

In the first of four studies we demonstrate anticipation of lexical items via contextual constraint, as well as noting an effect suggestive of an expectancy violation cost. In the second experiment, we demonstrate that predictive processing and its potential costs generalize, to a large extent, across input rate, offering support for fast, unconscious preactivation of linguistic content. In the third study, we test the nature of our proposed prediction cost, ultimately linking it to graded, general constraint violation. In the final experiment, we show that though both cerebral hemispheres are sensitive to message-level constraint, the left hemisphere appears biased toward more predictive sentence processing. Throughout, we take multiple approaches to defining constraint and expectancy, and argue that traditional quantifications of these concepts may conflate rich sources of variability.

CHAPTER 1.
INTRODUCTION AND BACKGROUND

1.1. A general argument for prediction

Stop for a moment. Try to quiet the voice inside your head. This may be difficult, likely impossible. But chances are there are still low rumblings along the lines of '*what's next?*' For instance, you might be thinking *How long should I wait? Should I scratch the itch on my leg? What is that noise in the hall? Should I eat Chinese or Mexican for lunch? Will I ever finish this thesis/publication/book/report/grant application? Is that new person in the lab a potential friend? Is it going to rain today? Will I find a job this year? Will my children grow up to be kind? How long will I live?* We humans possess a remarkable tendency to think in advance, anticipating consciously, for purposes of preparation and simulation of events on near and far temporal horizons. Yet at the same time, a multitude of unconscious predictions may also be forming at a variety of levels. Your heart rate may increase as you prepare to address your colleagues in seminar. Your attention may shift when you hear voices in the next room. Your hand position may subtly transform as you reach to grasp the pen on your desk. Without thinking, you step wide over the extension cord strung across the room to avoid tripping. Some of these "unconscious" processes are attributable to biology; some to learned associations formed through our experiences with people, places, things and events in the world; and still others to some interaction of these two factors. On the whole, then, the brain appears to function both aware and unaware in an anticipatory mode, with prediction of probabilistic outcomes providing undeniable benefit in preparing for upcoming events and avoiding danger.

However, anticipation could also be considered a risky strategy if projected outcomes do not obtain, for instance when a predictive gamble results in a reduction of available physical or mental resources. For life-threatening situations, a loss of resources incurred through anticipation might seem to be outweighed by the general benefits afforded by such processing. For instance, the initiation of a fight-or-flight response in a rabbit that ultimately avoids detection by a preying coyote might expend a certain amount of energy, but the price seems a small one to pay. But what about less vital situations – ones where the risks of not anticipating are arguably lower? In particular, the question of interest for this thesis is whether anticipatory processing is evidenced for a higher cognitive process such as language comprehension. Language, though perhaps uniquely human, is nevertheless implemented in brains with neural mechanisms similar to "lower" species. Thus, it could be argued that anticipatory processing is potentially as feasible and adaptive for language comprehension as it is for other real-time information processing domains across the phylogenetic spectrum. In this thesis we will explore the possibility that even if prediction turns out not always to provide an advantage in comprehending a stream of linguistic input, the brain may nonetheless be committed to such an approach as a natural consequence of how meaning is stored and constructed.

1.2. A brief proposal for a predictive language comprehension mechanism

Let us pretend that we are designing a language comprehension system to be instantiated in a human brain, one that meets all the necessary physiological processing constraints of a neural system. The language input we encounter (for the most part in visual, auditory, and gestural modalities) is generally serial, though through multiple sources there may be some temporal overlap. Optimally, the system should be able to process the incoming language stream while taking into account not just the influence of the linguistic input itself, but also the circumstances of our immediate surroundings and stored representations relating to the content at hand. The external input (e.g., the mode, style, or delivery of the language stream; speaker information; physical surroundings; etc.), our own internal state (e.g., arousal level, mood, or hormonal state), and the intersection of these factors are likely to influence the particular set of representations in memory that are activated as the sensory signal moves along the processing stream. The sounds and sights that we encounter are processed at multiple levels, building up from syllables (in the auditory modality) to words to phrases to sentences to discourse to inferences. Any given combination of input (and internal state) has the potential to be unique, though components of such combinations are likely to be familiar. When an input component or combination is received that corresponds with a stored representation (an activated network), there is likely ancillary information that gets activated "for free" (e.g., information associated as a function of encoding events or circumstances that co-occurred in time or place; or, for instance, due to phonological, orthographic, categorical, or semantic organization of stored memories, to name a few). Some of this ancillary (i.e., activated but not encountered in the physical stimulus) information may be relevant for the immediate linguistic circumstance, though some will not be. As the language signal continues to stream in, the accruing context may serve to highly activate a particular association or schema such that possibilities for what might come next are limited within an activated framework: other input patterns may lead to less precise correspondence with a structured memory, possibly with fewer "ancillary bits" getting activated, or possibly at weaker levels. As context accrues, new information highlights the potential importance (or lack of importance) of the previously activated "ancillary bits" such that their activation levels are up- or down-graded accordingly. It is important to keep in mind that these "ancillary" activations may occur at a variety of different levels (e.g., featural, syntactic, thematic) and in fact are likely to vary greatly across individuals. Even combinations of linguistic input which have never before been encountered "in the wild" have the potential to activate predictable patterns of information due to potential associative processes. And importantly, the "for free" bits are activated outside of conscious processing, much as when you glimpse a visual scene depicted with tables, waiters, menus and wine, one does not have to *think* "restaurant" but rather just *knows* that the scene is depicting a restaurant. At any point within a word, phrase, sentence or discourse, then, there is information which we might call "expected" or "predicted" because, at one level, many or most comprehenders may have formed common associations with the input patterns, which is evidenced by individuals converging on similar continuations when asked to "fill in the blank" with the most likely completion in offline questionnaires. These so-called "cloze procedures" (Taylor, 1953) are one of the few ways we have – though not a very elegant one – of tapping into the organization of semantic memory, though ratings are generally averaged across individuals and answers are supplied outside the temporal constraints of natural language processing. However, another way to think about linguistic input (or features of input) being "predicted" is simply in terms of the degree to which it (they) has/have been triggered – in advance of encountering the input – through links in long term memory. This is not to say that *all* language processing takes place unconsciously. Nor does it mean that there are instances in which the input encountered does not correspond with what has been activated "for free". Language, indeed, is full of surprising continuations, some of the more obvious ones being joke punch lines, frame shifts, cross-language code switches, and non-sequiturs. But in general, such a system offers a benefit – if you want to call it that – because information that might be encountered in upcoming input has already been pre-activated. This "benefit" of eased target integration might be considered more as a "bonus", because it is actually a byproduct of how knowledge is stored and activated rather than reflecting a system that is driven predict. Such prediction, nonetheless, may offer a biological advantage in freeing up the resources that would otherwise be required to activate, access and integrate each new bit of linguistic information in a bottom-up fashion. And when the prediction process fails to preactivate information that is eventually encountered in the input stream – information that is more likely to, but need not, alter the course of the already-accrued contextual representation in some significant way – then it might be assumed that there needs to be additional processing to reanalyze the representation that has accrued up until that point.

We sketch this picture of linguistic prediction for a few reasons. First, because when considering why an anticipatory mechanism might be a feasible approach for language comprehension, it is important to let the organization of meaning in the brain inform our view. If one thinks of prediction solely in terms of the infinite combinations of words available for conveying any one idea, anticipatory processing admittedly does not make much sense. But in language, as in other neural domains, receipt of an input is unlikely to trigger an infinite number of possibilities; rather, there is *always* some context (even when by experimental design there is no or neutral context!) such that continuations are inevitably constrained to at least to some degree. Preactivations of the sort we are proposing need not reach consciousness – though with time, they sometimes may – and would facilitate comprehension in a majority, though not in all, cases. The second reason we propose a view of prediction from the outset is to head off potential arguments against predictive processing on grounds of misconstruing what we take "prediction" to mean. There has historically been opposition to (and silence on) the subject of anticipatory language comprehension, for reasons which we suggest in the following section. The tide about predictive language processing seems to have changed in recent years, but it is still informative to review from where traditional (and in some circles, unabated) resistance to prediction has emanated.

1.3. Opposition to prediction in comprehension models

Despite general considerations about biological continuity and informal intuitions stemming from our experiences "taking the words out of" a speaker's mouth, the concept of pre-activating linguistic information has played a relatively minor role in theories of language processing. The idea that individuals might be predicting linguistic features or content has not been part of the generative grammar tradition. Chomsky (1957) and others (e.g., Jackendoff, 2002) have argued that with infinite options available as each new word of an unfolding sentence comes in, predicting what comes next is improbable. Unlike the comparatively small, fixed repertoire of non-human animal communication schemes, human languages offer unlimited possibilities for communicating information – far too many ways, they contend, to make lexical prediction a viable strategy except on the rare occasions when contextual constraint is unusually high (Stanovich & West, 1979).

Nonetheless, a basic observation from controlled experiments is that when participants are asked to supply lexical completions for truncated phrases or incomplete sentences, their responses tend to converge on the same words when contextual constraint is strong (e.g., Bloom & Fischler, 1980; Schwanenflugel, 1986). In off-line language tasks, then, it is widely acknowledged that with *sufficient time* individuals are capable of using sentential context to select the most probable linguistic completions. This approach, however, does not necessarily extend to sentence contexts which are less constraining, nor does this more deliberate, conscious (post-lexical) strategy necessarily translate to the rapid, less conscious processing that seems to characterize real-time language comprehension.

To be clear, and as a guide to interpreting the phenomenon we will be examining in this thesis, although the terms "prediction" and "anticipation" will be used interchangeably throughout this dissertation, a better way of understanding the process of what we consider to be anticipation in on-line processing is in terms of "pre-activation"; in other words, through associations in long-term memory, a context might automatically activate the mental representation of items (e.g., words, pictures, or concepts) and/or their features. An outstanding question in the sentence processing literature has been whether information about particular words or features of words gets pre-activated during on-line sentence processing as a result of top-down contextual processing, or whether processing of an item is initiated only after the physical stimulus has been received – a more bottom-up view. Perhaps surprisingly, there has not been clear consensus on this issue. In fact, most theories of language processing have not addressed the idea directly, though some stance about linguistic prediction is often inherent in different theoretical views.

Understandably, resistance to predictive language comprehension models may stem from the lack of experimental evidence for predictive effects that has characterized the field until quite recently. Finding proof of an event that has not yet taken place is difficult and requires both clever experimental manipulation as well as measures and paradigms that are highly sensitive to the timing of on-line sentence processing. Although under certain experimental designs prediction during sentence processing might be *inferred* from a given pattern of effects observed *at* a word target, such effects are generally more open to interpretation and are not as likely to be construed as the concrete findings needed to convince those dubious of prediction. Such "post-lexical" effects have been debated on grounds that facilitated processing (e.g., decreased naming latencies and lexical decision time, increased priming effects, etc.) observed at lexical targets might be due to preactivation of an item or its features, but they might also be due to the eased integration afforded by the context *upon* an item's presentation. The prediction/integration debate – an issue to which we will return later - then, is clearly one of timing, and therefore one not well adjudicated by some of the earlier examinations of sentence processing employing offline techniques.

Longstanding arguments against anticipatory language comprehension have focused on several main points. The first is that natural language is not considered to be constraining enough for a predictive system to be accurate a majority of the time, and such errors should, in principle, result in processing costs; however, traditionally such costs have not been evidenced (Gough, Alford & Holley-Wilcox, 1981). Other arguments center on the idea that word recognition can benefit from effects of context, but only when target input is degraded, when targets are offset by a temporal delay from the context, or when readers are unskilled (Mitchell, 1982). Rayner, Ashby, Pollatsek & Reichle (2004) go on to point out that findings for predictive processing based on studies using non-naturalistic presentation (e.g., artificial timing of sentences or context-target delays) allow readers to form conscious predictions about targets – circumstances that do not map to normal reading. A more pervasive component to the downplay of a role for prediction may have to do with the long history of more modular views of language processing (e.g., Forster, 1989), and their inherently bottom-up biases. A combination of these reasons, and perhaps others, has likely contributed to the general lack of exploration of anticipatory comprehension.

1.4. Incremental comprehension models

Early models of sentence processing were proposed to include some form of memory buffer by which sentential elements were temporarily stored for later integration at phrasal, clausal or sentence boundaries (Kintsch & van Dijk, 1978; Carroll & Bever, 1978; Just & Carpenter, 1980; Daneman & Carpenter, 1983; Mitchell & Green, 1978; Abney 1989; Pritchett 1991). However, if buffering had been a valid model for sentence comprehension, then effects of context updating on a word-by-word basis should not have been observed.

Instead, there has been ample empirical support over the years for sentence processing models in which word comprehension is affected by the context in which words appear. Notions of buffering gradually gave way to the commonly held view that comprehenders incorporate words successively in a bottom-up fashion into the sentential context as they are received and identified – an *integrative* view (e.g., Marslen-Wilson, 1975; Altmann & Steedman, 1988; Pickering, 1994; Tyler & Marslen-Wilson, 1977; Traxler, Bybee & Pickering, 1997; Marslen-Wilson & Tyler, 1980; Steedman, 1989; Boland, Tanenhaus, Garnsey & Carlson, 1995; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Eberhard, Spivey-Knowlton, Sedivy & Tanenhaus, 1995; Kutas & Hillyard, 1983). Under integrative accounts, comprehenders are not building up predictions based on prior context as a sentence unfolds, but rather the processing is thought to initiate upon receipt of the physical input (e.g., Seidenberg, Tanenhaus, Leiman & Bienkowski, 1982; Forster, 1981; Gough, 1972).

The view that language comprehension is indeed at least incremental is rooted in evidence from a wide variety of studies and methodologies. For instance, Boland et al. (1995) argued for incremental processing on the basis of a behavioral task. Using a paradigm that manipulated filler-gap dependencies, participants were asked to indicate with a button press at which point a sentence stopped making sense. Linguistic constructs of fillers and gaps are present in 'Wh'-questions such as, '*Which preschool/military base did Hank deliver the machine guns to __ last week?*' *Which preschool/military base* is the filler and the gap occurs after *to* where the theme (or direct object) role would slot. Participants pressed the button at *machine guns* for the *preschool* version of the sentence but not for the *military base* version. These results suggest that the filler is assigned a thematic role as soon as the verb *deliver* is encountered, arguing against a buffer and in favor of more incremental processing. Marslen-Wilson (1975), using another behavioral technique, showed that so-called "fast shadowers" – individuals required to verbally repeat back recorded speech with as little delay as possible (sometimes as short as 250 ms) – corrected mispronunciation errors in the recorded speech signal, indicating that they were processing the shadowed text at a semantic level.

Another methodology used to advance incremental models has been eye-tracking. In eye-tracking studies of linguistic comprehension, the time-locked characteristics of eye movements provide information about the processes that underlie spoken language comprehension. Experimenters have frequently used the 'visual world paradigm' (Tanenhaus et al., 1995), a design in which individuals' eye movements are monitored as they look at a visual scene and simultaneously hear a sentence or set of instructions that refers to objects in that scene. It can then be observed at which point during the sentence individuals begin looking towards mentioned objects, with subsequent inferences being made about the time course of sentence processing. For instance, Eberhard et al. (1995) and Tanenhaus et al. (1995) found that Ss made saccadic eye movement to objects immediately after hearing relevant words in a set of verbal instructions. Again, this evidence supports integration on a word-by-word basis.

Yet another method for tracking the time course of language comprehension has been event-related brain potentials (ERPs). ERPs provide a series of snapshots of the synaptic potentials generated primarily by multiple pyramidal cells in the neocortex firing synchronously. Scalp recordings of the coordinated activity of these neurons, which are thought to perform the computations critical for comprehension and cognition, afford a continuous but time-stamped look at ongoing neural activity as comprehenders make sense of sensory and internal stimulation they encounter while comprehending language. As an example of how ERPs can be used to demonstrate more incremental processing, Kutas and Hillyard (1980) showed that a semantically anomalous word within a certain context elicits an enhanced ERP component known as the N400 at the point at which it is encountered, relative to a semantically congruent continuation in the same context. The N400 – as a neural response to any potentially meaningful item - is often taken as an index of the difficulty of semantic integration, and as such, provides additional evidence for context updating before the end of the utterance. These and other studies then, indicate that at the very least, sentence processing is incremental.

1.5. Anticipatory sentence comprehension

More current empirical evidence argues strongly for sentential constraint beginning to exert its influence before a word has been uniquely identified (e.g., Hoeks, Doedens & Stowe, 2000; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Federmeier & Kutas, 1999; Wicha, Moreno & Kutas, 2003a; Wicha, Bates, Moreno & Kutas, 2003b; Wicha, Moreno & Kutas, 2004; Altman & Steedman, 1988; MacDonald, Pearlmutter & Seidenberg, 1994; Kamide, Scheepers, Altmann & Crocker, 2002; Kamide, Altmann & Haywood, 2003; Otten, Nieuwland, & Van Berkum, 2007; Federmeier, 2007). While such evidence is relatively recent, a few early language processing models argued for more anticipatory processing that used available contextual factors to activate words in advance of receiving them (Elman, 1990; Elman & McClelland, 1986; McClelland & Elman, 1986). The *prediction* view of language comprehension, then, suggests that comprehenders utilize a variety of sources of constraint as they become available to preactivate upcoming items, with aspects of linguistic input being processed both incrementally and in parallel. Factors influencing the build-up of contextually anticipated material include various phonological, syntactic, semantic or pragmatic dependencies, and predictions can arise at these various levels of processing. For instance, many of these studies listed above have indicated that prediction need not be at the level of specific word forms, but could instead be at the level of features or meaning.

While some proposals for comprehenders anticipating upcoming sentential information have relied on off-line measures (e.g., Schwanenflugel & Lacount, 1988, using lexical decision times to argue that high constraint contexts increase readers expectations for what might follow), in general more recent studies of linguistic prediction have taken advantage of ERP and eye-tracking methodologies, which offer the exquisite temporal sensitivity necessary to track sentence processing both *at* and *prior to* encountering language-related targets. Such techniques are useful for addressing the *prediction/integration debate*, an argument whose adjudication is made difficult due to the challenge of distinguishing facilitation effects at a sentence target as being due to words being predicted versus words simply being easier to integrate upon their receipt. The N400 is one example of this interpretational confound. While the component's amplitude is sensitive to a variety of factors – including word frequency, repetition, and concreteness, among others – N400s are especially large to nouns that do not semantically fit with their preceding context (e.g., the word *dog* in '*I take my coffee with cream and dog*', Kutas & Hillyard, 1980). N400s are also evident in responses to all but the most highly expected of nouns, even when they do fit with a prior sentence context. But despite the sensitivity of the N400 to offline semantic expectancy, it is impossible to determine whether the variation in N400 amplitude to the eliciting word during online sentence processing means that readers are using context to generate expectancies for that upcoming item (*prediction view*) or whether they are forced by the word itself to devote more or less resources to integrating the word into the existing sentence representation (*integration view*). Clearly, an argument for information getting pre-activated would be strengthened if it could be demonstrated that predictions were being formulated prior to target words.

To this end, eye-tracking methods have successfully employed the visual world paradigm to note preferential looks to visual targets before hearing the actual words (e.g., Altmann & Kamide, 1999; Kamide, Altmann & Haywood, 2003; Kamide, 2008; Altmann & Kamide, 2007; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Sussman & Sedivy, 2003). Unlike the eye-tracking evidence used to argue for integration, looks to candidate entities in these studies *preceded* the relevant input. For instance, upon hearing a sentence fragment such as '*The girl will ride the…*' while viewing a scene depicting a man, a girl, a motorcycle and a carousel, comprehenders looked toward the depiction of the carousel during the verb *ride*; conversely, upon hearing '*The man will ride the…*', they looked toward the motorcycle during the verb *ride* (Kamide, Altmann & Haywood, 2003). Studies such as this one led to conclusions that focused on such phenomena as the pre-assignment of thematic roles as well as verb (or noun+verb) information placing selectional restrictions on upcoming input. These findings provide evidence that at least when candidate visual entities are present and easily accessible in the surrounding context, the language parser can act quickly to narrow the possibilities for upcoming input.

There has also been some ERP research that has argued inferentially for prediction, based on effects at, but not prior to, target nouns. For instance, Federmeier & Kutas (1999) demonstrated that in high – but not low – constraint sentence contexts, unexpected noun targets categorically related to highly expected endings were processed more similarly to the expected endings than were unrelated unexpected nouns. The researchers suggest this pattern results from the preactivation of the common features of the categorically related item to the expected noun.

Other ERP studies have reported pre-target brainwave effects that indicate features of words (or the words themselves) have been preactivated (e.g., DeLong, Urbach & Kutas, 2005 - Experiment 1 of this thesis; van Berkum et al., 2005; Wicha et al. 2003a and 2004; Wicha et al., 2003b). In general, these studies have relied on clever experimental manipulations that allow researchers to examine ERPs for more and less expected sentence continuations at points temporally preceding semantic targets – points at which, in principle, there should be no difference in the ERP patterns unless an upcoming item was predicted. These manipulations have utilized various forms of prenominal marking that have little inherent semantic value themselves – e.g., grammatical gender marking and phonologically biasing determiners – but were associated with more semantically meaningful targets. For instance, van Berkum et al. (2005) found ERP differences between gender-marked prenominal adjective endings in Dutch that either agreed with or did not agree with highly expected noun continuations (e.g., *The burglar had no trouble locating the secret family safe. Of course it was situated behind a big*$_{\text{neuter/common gender}}$ *but unobtrusive painting*$_{\text{neuter}}$*/bookcase*$_{\text{common}}$*.*) These results suggest that comprehenders already had a particular (gender-marked) noun in mind well before its receipt.

Taken together then, these eye-tracking and ERP studies, which will be elaborated upon in subsequent chapters, begin to outline a picture of an actively predictive neural language parser, preactivating word features and forms in advance of input. Our hope is that the studies presented in this thesis will begin to "fill in" this outline, so that anticipatory processing might no longer be considered a lingering question in the literature, but rather be understood as a natural part of the way language is comprehended, with subsequent investigations probing the nature and consequences of prediction.

1.6. Alternative conceptualizations of prediction

So far, we have attempted to frame a brief proposal of how we believe linguistic prediction might be instantiated in the brain, and we have also outlined a few of the methods in which the theory has been tested to date. However, at the same time that much of this experimental work has been taking place, there have been parallel explorations of this topic by computational modelers. Many of the models of human sentence processing have focused on anticipation in terms of syntactic rather than semantic (or even more integrated) forms of prediction. For instance, Levy (2008) frames an expectancy-based theory of syntactic comprehension in terms of *surprisal* values, where the difficulty of a word in context relates to the conditional probability of an item. Jurafsky (1996) also presents a probabilistic model for lexical access and syntactic disambiguation, but again one which is most concerned with syntactic processing. Gibson's (1998) syntactic prediction locality theory has an explicit anticipatory component, in which syntactic requirements are initiated upon receipt of an input and unfulfilled predictions tax processing resources. Other models, such as McRae, Spivey-Knowlton & Tanenhaus (1998), integrate syntactic, lexical and semantic constraints in their networks in order to predict features of upcoming words in sentences. Still others have aimed to integrate visual scene information with linguistic experience in their comprehension models (e.g., Crocker, Knoeferle & Mayberry, 2009).

We believe that attacking the prediction problem from multiple fronts can only lead to richer theories and more testable hypotheses, but in our view the models that consider multiple sources of constraint and argue for prediction not just at the syntactic level are likely to be the most promising. This is important because the kind of prediction that we are proposing is multifaceted, likely to occur not just for syntactic elements (parts of speech, thematic roles, morphological endings, etc.), semantic continuations (words, pictures, signs, gestures, etc.), and phonological characteristics (e.g., word-initial sounds, rhyme schemes, etc.), but also for subtle features and affordances of linguistic items, which we propose may get triggered through activation of various experience-based networks in semantic memory. Ideally, both modeling and experimental approaches will inform each other, with experimental evidence generating phenomena to model, and models uncovering potentially new applications. In terms of costs/benefits, constraints/informativity, and surprisal/processing time tradeoffs, computational models offer a complementary route for investigating anticipatory language processing. However, we believe that generating a set of solid experimental prediction findings has been, and continues to be, a major avenue for advancing the theory.

Similar to a generally syntactic emphasis in the models of linguistic prediction, there is also a bias towards more syntactic forms of anticipatory processing in experimental based theories. Some language comprehension frameworks allow that contextual constraint could theoretically be guiding predictions of syntactic components such as verb arguments; however, under proposals suggesting that semantic information does not get stored with the verb information, constraint may be doing little to shape predictions of semantic properties of those upcoming arguments (e.g., Pollard & Sag, 1994). Other theories propose that semantic information of verb arguments may not get preactivated because access to such information is delayed until after syntactic integration takes place (Rayner, Carlson & Frazier, 1983). These theories are less compatible with our notion of prediction than theories suggesting, for instance, that the semantic properties of verb arguments are projected upon receipt of the verb, and importantly, prior to encountering the noun phrase in the linguistic input (e.g., Kamide, Altmann & Haywood, 2003).

Thus it is worth delineating our view of what anticipatory language prediction may entail from other proposals of what constitutes prediction in the sentence processing literature. For instance, language predictability is a topic that has been discussed frequently in eye-tracking studies of reading. Predictability, in reference to eye movements to particular items, has often been meant to refer to the findings that words considered more 'predictable'- i.e., highly constrained in a certain context - are directly fixated less frequently during reading than unconstrained words (Balota, Pollatsek & Rayner, 1985; Ehrlich & Rayner, 1981). They are also regressed to less than unconstrained words (Ehrlich &

Rayner, 1981; Inhoff, 1984) and show shorter fixation times than unconstrained words (Altarriba, Kroll, Sholl & Rayner, 1996). Such effects, though, in our view, do not deal directly with the prediction question as we have outlined it herein. These effects do not address the timing issue – that of noting evidence for items being predicted prior to encountering them – and again are open to interpretation in terms of both "integrative ease/difficulty" as well as prediction.

1.7. The brain as a global predictor

Altmann & Kamide (2007) state that '*Anticipatory eye movements do not reflect the unfolding language; they reflect an unfolding (mental) world*.' Here, Altmann & Kamide refer to the insights offered by a specific methodology into brain processes that function predictively in general meaning construction. This view, and the one we explore in this thesis in relation to anticipatory language comprehension, though, are compatible with a wave of recent treatises on how neural processing across domains and as an organizational principle is predictive. In particular, early explorations of perception as being inherently predictive in nature have expanded to include theories about higher level cognitive processes being "proactive" (Bar, 2007). Indeed, there have been recent theme issues ('Predictions in the brain: using our past to prepare for the future', *Philosophical Transactions of the Royal Society, Biological Sciences: May 12, 2009*) and 'trends' articles ('Do people use language production to make predictions during comprehension?', Pickering and Garrod, 2007; 'Prediction of external events with our motor system: towards a new framework', Schubotz, 2007 – both from *Trends in Cognitive Science*) devoted to exploring prediction in various domains. Investigations run the gamut across topics as diverse as cortical versus subcortical predictive brain dynamics, the role of prediction in emotional processing, prediction in motor control, predictive mental imagery, prediction and self awareness, the cellular machinery of prediction, and the role of prediction in decision and judgment making. A common thread across many of these articles seems to be the role of memory in predictive processing, which fits nicely with our current investigations of prediction in language comprehension. Though this dissertation will focus on anticipatory brain processing within one particular domain, the accruing evidence for general memory-based predictive brain processing at the very least frames the proposals we present herein as being consistent with other neural processes. However, further inter-domain comparisons – while intriguing – will be beyond the scope of our current investigations.

1.8. Consequences of prediction

A flip side to the idea that comprehenders might be preactivating anticipated upcoming information during online language processing, is that there might be processing costs when the language parser has to revise a strong prediction when a sentence continuation is unexpected. Most comprehension theories that include processing costs posit them in terms of syntactic processing. For instance, a theoretical framework with a prediction component – the Dependency Locality Theory (Gibson, 1998) – models grammatical complexity in sentences in terms of integration costs as a consequence of predictions generated from previous items, with such costs being influenced by the distance of syntactically dependent elements. Other studies have proposed that the P600, a late occurring positive-going ERP component, indexes some form of syntactic processing cost, though its exact nature has been debated as alternately reflecting syntactic disambiguation (e.g., Osterhout & Mobley, 1995), revision (Osterhout & Holcomb, 1992), integration difficulty (Kaan, Harris, Gibson, Holcomb, 2000), or reanalysis (e.g., Hagoort, Brown & Groothusen, 1993). Syntactic processing costs seem to have gained relevance in the sentence processing literature due to the role they may potentially play in differentiating serial and parallel processing. For example, with parallel processing models, costs might be proposed because alternative interpretations of ambiguous input are simultaneously pursued; conversely, the idea of cost would not necessarily come into play for a serial processor, in which a single structural analysis is pursued for a syntactically ambiguous item.

So what, then, would constitute a cost for semantic processing during sentence comprehension? In early word priming paradigms, inhibition was sometimes cited as a cost upon receiving semantically unrelated targets, though such effects were debated to dissipate at shorter interstimulus time intervals (Neely, 1991). Because prediction has, until recently, played a relatively trivial role in language comprehension theories, "cost" is a term that typically has not been referred to in conjunction with semantic processing. Rather, comprehension has typically been described in terms of relative ease or difficulty of integration. In the ERP literature, some have tried to identify the N400 component as reflecting a cost for semantic/conceptual processing, though we would argue that N400 parameters do not vary with cost, but rather with degree of facilitation (or the benefit) of a supportive context. After all, the N400 is the brain's default ERP response to a meaningful item and shows a *reduction* in amplitude as items become more constrained. However if one assumes a constraint-based comprehension system that is continuously preactivating, then a cost (e.g., for overriding, revising, inhibiting, or reanalyzing) – particularly in cases where there is a strong candidate that does not materialize – seems a reasonable subject of investigation.

Recently, there have been several studies which have proposed costs more related to semantic processing – or at least to the intersection between syntactic and semantic processing. Though, as mentioned above, the P600 ERP component has most typically been observed in response to specifically syntactic sentence manipulations, various research groups have observed a similar brainwave response to experimental manipulations that might be considered more semantic. Such "semantic P600s" have alternately been tied to experimental findings reflecting thematic role assignment costs (e.g., Hoeks, Stowe & Doedens, 2004), detections of conflicts between semantic plausibility and syntactic requirements (e.g., Van Herten, Chwilla & Kolk, 2006), and conflicting processing streams - including ones for syntax and semantics – (e.g., Kuperberg, 2007). None of the preceding proposals, however, directly implicates violation of general linguistic prediction as the possible source for such ERP effects, though a different study does. Federmeier, Wlotko, De Ochoa-Dewald & Kutas (2007) observed a late positive ERP effect to low probability congruous sentence endings that continued highly but not weakly constraining contexts, though this effect was not present during a lateralized presentation of identical stimuli (Wlotko & Federmeier, 2007). Taken together, these studies suggest that there may be some basis for attempting to isolate a cost for misprediction.

1.9. A hemispheric role for prediction

In recent work on neural prediction, there have been some proposals for more domain-general cortical networks associated with anticipatory processing and also with detecting errors in prediction (see Bar, 2007). However, with respect to predictive processing in the realm of language comprehension, there has been relatively little work that has implicated a role for specific brain areas. While the traditional view of language processing has been that both comprehension and production are handled primarily by the brain's left cerebral hemisphere, more modern views posit that at least some aspects of language processing - including discourse comprehension; nonliteral language processes such as humor, sarcasm, metaphors and indirect requests; and sensitivity to emotional prosody – are handled by the right hemisphere. Still other views indicate involvement of both hemispheres across a range of language processes spanning from word to sentence to discourse comprehension, albeit with different sensitivities, timing, or processing mechanisms.

As referred to above, Federmeier and colleagues have been one of the few groups to explore prediction-related language processing with respect to the role of the two cerebral hemispheres (e.g., Federmeier & Kutas, 1999b; Wlotko & Federmeier, 2007; Federmeier, 2007). Their general proposal has been that the left hemisphere exploits context in a more top-down fashion to preactivate information in semantic memory that may eventually be encountered in sentential input; conversely, the right hemisphere functions in a more integrative fashion, linking up linguistic content with existing contextual representations at the point the input is received. Under this proposal, one might also postulate that prediction errors (at contextual constraint violations) would also be preferentially handled by the left cerebral hemisphere, though to date such evidence has remained elusive (Wlotko & Federmeier, 2007).

In general, a hemispheric division in which one half of the brain anticipatorily activates linguistic content in a probabilistic manner while the other half engages in more bottom-up processing, might offer a sort of check-and-balance system to prediction gone wrong and a mechanism by which recovery from such instances might be facilitated. Though from a design perspective such an organizational principle makes sense, the validity of such a model must be followed up on and substantiated empirically, particularly with regard to the consequences of mispredicting. If indeed, there is a hemispheric bias in the sensitivity to violations of contextual constraint, then a number of questions are spurred. For instance, is the nature of a prediction "cost" graded or all-or-none? Are the hemispheric divisions in predictions similar to those for recovering from misprediction? And are the timings of such processes equivalent across the brain's two halves?

1.10. Event-related brain potentials (ERPs) for studying linguistic prediction

1.10.1. The ERP technique

As mentioned above, some of the past hesitancy of the sentence processing community in adopting more predictive models of language processing may be due in part to the practical difficulties of designing experiments that allow anticipation to be tracked during language comprehension. Observing anticipatory effects in sentence processing requires a measure that has temporal resolution on the order of milliseconds and does not alter the comprehension process under study. Also necessary is a methodology that allows continuous monitoring of responses throughout the course of a sentence so that observations can be made not just at target words, but also at words preceding targets. For these reasons, and others, the primary methodology that will be used to track prediction in the studies of this thesis is the event-related brain potential (ERP) technique.

ERPs are the electrical brain activity synchronized in time to a stimulus, response, event, or even an event's absence. Recordings, like the ones performed in the studies of this thesis, are collected at the scalp with surface electrodes embedded in an electrode cap or net. ERPs are presumed to be the brain's (or at least the cortical regions') response to eliciting events, comprised of millisecond-by-millisecond time series of voltages at individual recording sites, affording inferences about brain sensitivities and computations. Because the time-locked response to an event coincides with electrical brain activity not time-locked to the eliciting event (either at all or in phase), ERPs are typically extracted via averaging across multiple repetitions of the eliciting event. These repetitions need not be exactly the same stimulus or event as long as they are conceptually similar with respect to the experimental manipulations: for instance, averaging the ERPs within a condition comprised of forty different, but similarly anomalous, sentences. The ERP waveform is used to track the course of information processing as "sensory inputs" travel from the eye or ear to the mind for comprehension, or in reverse, as "outputs" travel from mind to tongue or hand, for production. The scalp ERP, at each recording site, is an instantaneous reflection of the sum of all the brain activity that meets the criteria for being recordable at the scalp surface (see Kutas & Dale, 1997). An oft-cited benefit of this method is that brainwaves are elicited without need for any overt motor response, a particular advantage when measuring how such signals change, for instance, over normal or abnormal development, as a function of healthy or pathological aging, or due to various sorts of brain damage.

ERPs offer an image of cortical brain activity across time at multiple timescales, from milliseconds to seconds in a continuous and relatively non-invasive manner. The moment of stimulus onset is considered "zero" on the time axis, and in this way the event related brain potential is temporally synchronized to the eliciting event. With respect to this "zero" time point, however, what exactly the event-related potential voltage waveforms represent is a matter of some theoretical debate. On one view, it is assumed that the electrical brain activity synchronized to stimulus onset is a reflection of the brain processing of that item at that particular point in time. On another view, the presentation of the stimulus item acts to perturb (and re-align) the ongoing EEG rhythms. In either case, such recordings provide a time series of dependent measures generated by the brain's processing of the events that can be meaningfully interpreted in the context of any experimental design to reveal something about how the brain construes or produces language inputs.

While the methodology has many attributes to recommend it for the study of cognitive phenomena, the "reading" of the actual ERP waveforms is neither straightforward nor intuitive. Although electrical brain activity is sensitive to and reflects a wide variety of factors (including how recently one ate, how tired one is, or a person's general mood at the time of the experiment), the elicited potentials do not come with a roadmap. At any given point, the ERP waveform is negative, positive or neutral relative to some baseline, but these negativities and positivities per se have no intrinsic meaning relative to their neural generators. Nonetheless, cognitive ERPers have been developing a glossary of ERP components and effects by discovering manipulations that usually lead to them. These in turn are defined in terms of their polarity, shape, topography across the scalp, and latency relative to zero. By convention, the labeled patterns get their names either from a combination of their polarity and/or approximate latency (e.g., N400 or P600) or by their theorized functional significance (e.g., Syntactic Positive Shift or SPS). These components then serve as proxies for particular constructs or information processing operations (e.g., semantic processing, working memory updating, grammatical processing, etc.) and can sometimes be used – as we will do in this dissertation – to adjudicate between alternative theoretical accounts of various psychological phenomena.

1.10. 2. ERPs and language processing

Less than half a century ago, the idea of investigating the psychology of language by recording electrical activity from the human scalp seemed a fantastical notion, like a concept lifted from the pages of a science fiction novel. By the 1980's, however, the idea of using electrophysiology to study linguistic phenomena was becoming a respectable enterprise. Different types of electrical brain responses were being observed at the scalp to words that were incongruent with prior contexts at semantic, syntactic, and phonological levels. It had also been shown that ERPs were sensitive to psycholinguistic variables of all sorts in perfectly normal sentences, even when all of the experimental sentences were grammatically well formed and meaningful. Such findings paved the way for the virtual explosion of ERP and magnetoencephalographic (MEG) investigations of language processing over the past quarter century. With diverse studies running the gamut from phoneme categorization to discourse processing, with groups ranging from healthy to brain damaged individuals, with measures recorded from populations spanning infancy through old age, virtually no area remained untouched by cognitive ERP researchers.

In many cases, ERP effects are highly correlated with overt behaviors and lead to inferences that could just as easily be made from external performance measures such as various speed and accuracy judgments, or the scanning patterns of eye movements across either scenes as people listen to sentences or across printed words as they read text. Such converging measures are important for amassing databases of psycholinguistic phenomena and for constructing comprehensive theories. However, ERP and performance measures are at times dissociable and can occasionally offer different pictures of the same cognitive acts; thus, for thorough understanding of a particular area of interest, the different perspectives afforded by the varying methods must be integrated. Perhaps most importantly though, and in response to those who might advocate the use of only more basic behavioral methods for reasons of time and expense, there are also cases in which ERPs provide a unique view into moment-by-moment cognitive processing, with no parallel in other measures. While the limitations of various performance measures are sometimes clear at the point when the experimental design is conceived, in other instances it is not possible to tell whether more basic methods will provide sufficient power to reveal potential differences between conditions. As experimental design does not come with a crystal ball, it is often only with subsequent testing using more neurally-informed and informative methodologies, such as ERPs, that differences between conditions are revealed where none were originally detected.

Because so much of language processing occurs quickly and is unavailable to conscious reflection, ERPs have made some of their greatest contributions in probing natural language processing. Prediction, in particular, is one phenomenon that is not so easy to capture, the difficulty being in how to observe the consequences of an event that has not yet occurred. For this reason, ERPs have been a methodology of choice. The electroencephalogram can provide a continuous monitor of brain responses throughout the course of a sentence or discourse, i.e., prior to target events that might be subject to prediction. ERPs can also provide a measure of the quantitative and qualitative changes that may distinguish two conditions, including sensitivity to semantic and syntactic variables and processing. In these ways, ERPs can be used to explore research questions that behavioral or in some cases even eye-tracking studies cannot as easily resolve.

1.11. Primary theoretical questions of thesis

This thesis will utilize event related brain potentials (ERPs) as the dependent measure in identifying and elaborating upon the nature, preconditions, consequences, and neural processes underlying anticipation in on-line language comprehension. We will attempt to establish clear evidence for whether preactivation of lexical word forms occurs in the course of comprehending written language and in turn, determine whether such prediction incurs a processing cost when anticipated word forms do not occur. We will probe the consistency of these potential findings across input rates and over a range of more and less constraining contexts, as well as examine the cerebral hemispheric contributions to such processing. In total, four ERP experiments will be presented, each examining some aspect of brain responses potentially linked to preactivation during online sentence comprehension. We aim to clearly dissociate predictive from integrative comprehension effects, as well as investigate the probabilistic nature of constraint-based prediction and its consequences. In this way, we hope to determine whether preactivation and the potential associated costs of incorrectly predicting are general language comprehension mechanisms, or ones that engage only under a limited set of circumstances. Specifically, this thesis will address the following four questions.

1.11.1. Is there contextual constraint-based preactivation during online language processing and if so, what is its nature?

Experiment 1 is designed to distinguish between the hypothesis that specific phonological word forms are anticipated during real-time comprehension and the hypothesis that words are simply integrated with on-going contextual representations at the point at which they are encountered. Using a paradigm that allows us to monitor for lexical activation prior to a target word's presentation, we investigate such effects across sentences of varying constraint. The preactivation question is relevant because to date, strong evidence has not been obtained for prediction of specific lexemes during sentence processing, with even less evidence for linguistic prediction of a probabilistic nature. Our findings will be framed in terms of information activation through semantic memory.

1.11.2. What are some of the pre-conditions of prediction?

One factor that could potentially influence the nature of prediction effects is rate of input. Though results from ERP studies of sentence processing that present stimulus items at somewhat slowed reading or speech rates are frequently interpreted as generalizing to more natural language input rates, the issue of timing is one that seems particularly relevant to studies of linguistic prediction. The core question of Experiment 2 is whether effects of linguistic preactivation are a product of fast, unconscious processing or whether sentence information can only be activated through more conscious, deliberate strategies. This experiment utilizes visual sentence presentation and a stimulus set identical to that used in Experiment 1, except that rate of presentation is increased to one on par with more naturalistic reading rates. Thus, this study constitutes a first step toward testing whether the brain has a limited capacity to pre-activate information based on the timing of the input.

1.11.3. What are the consequences of anticipating incorrectly?

Traditional arguments against predictive mechanisms of language comprehension are that the potential to predict incorrectly is great and that there may be consequences to pre-activating information that never occurs. Prediction could be considered beneficial when accurate, but there might be a cost to expectations being partially or wholly disconfirmed. Chapter 4 (Experiments 3A,B) focuses on whether there are consequences to mispredicting and what the nature of such prediction costs might be. In particular, the costs might be presumed to be greater when a highly – but not weakly – constrained item is not received. We investigate these issues by first probing how comprehenders recover from the presentation of unexpected function words (grammatical determiners) in a sentence norming study, and then use this information as the basis for distinguishing between more syntactic versus more general prediction costs in an online sentence comprehension study. Based on our results, we then formulate a view of prediction costs in terms of constraint violation.

1.11. 4. Is there a hemispheric bias for linguistic prediction and its potential consequences?

Though the right cerebral hemisphere (RH) is no longer considered (as it once was) to be completely uninvolved in language comprehension processes, its role in extracting meaning from input and constructing representations during online sentence processing is still widely debated. Under one theory (Federmeier and Kutas, 1999; Federmeier, 2007) the left hemisphere (LH) is considered to function more predictively, utilizing sentential context in a more top-down manner to preactivate potential upcoming information; conversely, under this theory, the RH is thought to be more integrative, engaging in more bottom-up processing. If this distinction is accurate, then one might expect the LH to also be more sensitive to the consequences of violating sentential constraint. We test this proposal in Experiment 4 by using the visual hemi-field paradigm to isolate the contributions of the two hemispheres in comprehending sentences with a range of constraint and over a range of more to less expected continuations. The electrophysiological modulations of constraint and cloze probability on the ERP are described and contrasted between hemispheres.

1.11. 5. Summary of goals

Taken together, these four ERP studies are intended to offer some insights into how an anticipatory neural mechanism that is essential for survival across a variety of cognitive domains may also play a critical role in how the brain processes yet another set of environmental cues to meaning – namely, linguistic input. The evidence described herein for preactivation and its consequences during online sentence comprehension will demonstrate that language processing is not just incremental in nature, but rather incremental *and* anticipatory. These examinations of instances when preactivation works beneficially to facilitate meaning integration and when it leads comprehenders down the "wrong" path can ultimately inform our views not only of how contextual information taps into on-the-fly meaning construction and semantic organization in the brain, but also how brains (or their subsystems) function to override probabilistic projections.

1.12. References

Abney, S.P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research. Special Issue: Sentence Processing*, 18(1), 129-144.

Altarriba, J., Kroll, J.F., Sholl, A. & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, 24(4), 477-492.

Altmann, G.T.M. & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502-518.

Altmann, G.T.M. & Steedman, M.J. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

Altmann, G.T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Balota, D.A., Pollatsek, A. & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17(3), 364-390.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280-289.

Boland, J.E., Tanenhaus, M.K., Garnsey, S.M. & Carlson, G.N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34, 774-806.

Carroll, J. & Bever, T. (1978). The perception of relations: a case study in the relation of knowledge and perception. In: Carterette, E., Friedman, M. (Eds.), *Handbook of* perception, Vol. 7, Speech and language. Academic Press, New York.

Chomsky, N. (1957). *Syntactic structures*. Oxford, England: Mouton.

Crocker, M.W., Knoeferle, P. & Mayberry, M.R. (In press). Situated sentence processing: The coordinated interplay account and a neurobehavioral model. *Brain and Language*.

Daneman, M. & Carpenter, P.A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 9, 561-584.

DeLong, K.A., Urbach, T.P. & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117-1121.

Eberhard, K.M., Spivey-Knowlton, M.J., Sedivy, J.C. & Tanenhaus, M.K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409-436.

Ehrlich, S.F. & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior*, 20(6), 641-655.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science: A Multidisciplinary Journal*, 14(2), 179-211.

Elman, J.L. & McClelland, J.L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27(2), 143-165.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505.

Federmeier, K.D. & Kutas, M. (1999). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Federmeier, K.D. & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, 8(3), 373-392.

Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E. & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research. Special Issue: Mysteries of Meaning,* 1146, 75-84.

Forster, K.I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. (75-107). Cambridge, MA, US: The MIT Press.

Forster, K.I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, 33, 465-495.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.

Gough, P.B. (1972). One second of reading. In Kavanaugh, J. F., & Mattingly, I. G. (Eds.). *Language by ear and eye*. Cambridge: MIT Press.

Gough, P. B., Alford, J. A. & Holley-Wilcox, P. (1981). Words and contexts. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (85-102). Hillsdale, NJ: Erlbaum.

Hagoort, P., Brown, C. & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes.Special Issue: Event-Related Brain Potentials in the Study of Language*, 8(4), 439-483.

Hoeks, J.C.J., Stowe, L.A. & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19(1), 59-73.

Hoeks, J., Doedens, G. & Stowe, L. (2000). Lexical Prediction and Semantic Relatedness Interact in Sentence Processing: Evidence From ERPs. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, 30 March - 1 April 2000.

Inhoff, A.W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning & Verbal Behavior*, 23(5), 612-624.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic acces and disambiguation. *Cognitive Science: A Multidisciplinary Journal*, 20(2), 137-194.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.

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

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, Volume 2, Issue 4, 647-670.

Kamide, Y., Altmann, G.T.M. & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133-156.

Kamide, Y., Scheepers, C. & Altmann, G.T.M. (2003). Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, Vol. 32, No. 1, 37-55.

Kamide, Y., Scheepers, C., Altmann, G.T.M. & Crocker, M.W. (2002). Integration of syntactic and semantic information in predictive processing: Anticipatory eye-movements in German. Presented at the 14th Annual CUNY Conference on Human Sentence Processing, New York, USA, in March 2002.

Kintsch, W. & van Dijk, T.A. (1978). Towards a model of text comprehension and productions. *Psychological Review*, 85, 363-394.

Knoeferle, P., Crocker, M.W., Scheepers, C. & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, 95(1), 95-127.

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

Kutas, M. & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539-550.

Kutas, M. & Dale, A. (1997). Electrical and magnetic readings of mental functions. In M. D. Rugg (Ed.), *Cognitive Neuroscience*. (197-242). Cambridge, MA, US: The MIT Press.

Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126-1177.

MacDonald, M.S., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*. 101(4), 676-703.

Marslen-Wilson, W.D. & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

Marslen-Wilson, W.D. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226-228.

McClelland, J.L. & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.

McRae, K., Spivey-Knowlton, M. J. & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38(3), 283-312.

Mitchell, D.C. & Green, D.W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, 30, 609-636.

Mitchell, D.C. (1982). *The process of reading: a cognitive analysis of fluent reading and learning to read*. Published Chichester; New York: Wiley, 1982.

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. (264-336). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

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

Osterhout, L. & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34(6), 739-773.

Otten, M., Nieuwland, M. & Van Berkum, J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, 8(1), 89.

Pickering, M.J. (1994). Processing local and unbounded dependencies: A unified account. *Journal of Psycholinguistic Research*, 23, 323-352.

Pollard, C. & Sag, I.A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Pritchett, B.L. (1991). Head position and parsing ambiguity. *Journal of Psycholinguistic Research. Special Issue: Sentence Processing*, 20(3), 251-270.

Rayner, K., Ashby, J., Pollatsek, A. & Reichle, E.D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 720-732.

Rayner, K., Carlson, M. & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning & Verbal Behavior*, 22(3), 358-374.

Schwanenflugel, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instruments & Computers*, 18(4), 363-371.

Schwanenflugel, P.J. & LaCount, K.L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology*: Learning, Memory, and Cognition, 14(2), 344-354.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M. & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.

Stanovich, K.E. & West, R.F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition*, 7(2), 77-85.

Steedman, M. (1989). Grammar, interpretation, and processing from the lexicon. In: Marslen-Wilson, W. (Ed.), *Lexical Representation and Process*. MIT Press, Cambridge, MA.

Sussman, R.S. & Sedivy, J.C. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, 18(2), 143-161.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632-1634.

Traxler, M., Bybee, M. & Pickering, M. (1997). Influence of connectives on language comprehension: eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology*. A, Vol 50A(3), 481-497.

Tyler, L.K. & Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 683-692.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

van Herten, M., Chwilla, D.J. & Kolk, H.H.J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, 18(7), 1181-1197.

Wicha, N.Y., Moreno, E.M. & Kutas, M. (2003a). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483-508.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003b). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

Wlotko, E.W. & Federmeier, K.D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001-3014.

Portions of Chapter 1 (primarily descriptions of the ERP methodology) were extracted from a chapter which I co-authored along with Marta Kutas and Michael Kiang, to be included in the *Handbook of Psycholinguistic & Cognitive processes: Perspectives in Communication Disorders* (Eds. Jackie Guendouzi, Filip Loncke & Mandy J. Williams, Psychology Press, in press).

CHAPTER 2.
EXPERIMENT 1: PROBABILISTIC WORD PRE-ACTIVATION DURING LANGUAGE COMPREHENSION INFERRED FROM ELECTRICAL BRAIN ACTIVITY

2.1. Abstract

Despite the numerous examples of anticipatory cognitive processes at micro and macro levels throughout the natural world, the idea that anticipation of specific words plays an integral role in real time language processing has been a contentious one. Here we exploited a phonological regularity of English indefinite articles – *an*'s precede nouns beginning with vowel sounds whereas *a*'s precede nouns beginning with consonant sounds – in combination with event-related brain potential (ERP) recordings from the human scalp to show that readers' brains can pre-activate individual words in a graded fashion to a degree that can be estimated from the probability that each is given as a continuation for a sentence fragment offline. These findings are evidence that readers use the words in a sentence (as cues to their world knowledge) to estimate relative likelihoods for upcoming words. We additionally report on an unanticipated ERP finding to less expected sentence continuations, which we suggest may relate to processing of some type of expectancy violation.

2.2. Introduction

In a variety of everyday situations humans must react quickly to environmental cues in order to avoid harm or execute successful behavioral responses. If obstacles block our path, we want to avoid them before impact occurs, and when grasping an object we make slight (unconscious) corrections to hand position that account for the object's shape, distance, and direction. At the cellular level, nerve cells cause adrenaline release in preparation for fight or flight, and with learning, conditioned stimuli trigger brain responses that signal prediction of reward. From an evolutionary standpoint, these forms of anticipatory processing seem to enhance fitness by allowing for a series of small "advance decisions" that increase readiness and free up resources at more critical moments. Across a variety of physiological domains, then, the brain and body seem to have evolved in tandem as an "expectation machine." But the question of whether anticipation is inherent in a high-level cognitive process like language comprehension – which in many cases requires no response other than passive listening or reading – is an open one.

In terms of language comprehension, it is important to keep in mind what it is we mean by "prediction". Anecdotally, most of us have experienced "taking the words out of someone's mouth". And in off-line language tasks it is widely acknowledged that with sufficient time (under controlled processing) individuals are capable of using sentential context to select the most likely linguistic completions. In experiments when participants are asked to supply lexical completions for truncated phrases or sentences, their responses tend to converge on the same words when contextual constraint is strong (e.g. Bloom & Fischler, 1980; Schwanenflugel, 1986). However, the average comprehender of spoken English hears approximately three words per second – a relatively rapid rate of input, frequently beyond the listener's control. The debate, then, about anticipatory language processing seems to revolve not around what happens when there is a conversational pause or unlimited time to "fill in the blank", but rather around whether there is automatic pre-activation of upcoming linguistic items (or features of items) in the course of rapid, on-line comprehension.

In spite of the variety of real-time processing domains across the phylogenetic spectrum in which anticipatory processing has been observed and in spite of our own intuitions about anticipating, the concept of anticipation has played a relatively minor role in language processing theories. Human languages offer unlimited possibilities not only for saying new things but also for saying old things in new ways – far too many ways, some have argued, to make prediction of words a viable and effective strategy except when contextual constraint is unusually high (Stanovich & West, 1979). Accordingly, early language processing models often included some form of memory buffer wherein sentential elements were temporarily stored for later integration at phrasal, clausal or sentence boundaries (Kintsch & van Dijk, 1978; Carroll & Bever, 1978; Daneman & Carpenter, 1983; Mitchell & Green, 1978). Since the 1970s however, the consensus view has been that sentence processing is continuous and incremental, with provisional commitments made that at least temporarily resolve linguistic ambiguities as each word is processed upon its occurrence and rapidly integrated into the sentence representation (Marslen-Wilson, 1975; Altmann & Steedman, 1988; Pickering, 1994; Tyler & Marslen-Wilson, 1977; Traxler, Bybee & Pickering, 1997; Marslen-Wilson & Tyler, 1980; Steedman, 1989; Boland, Tanenhaus, Garnsey & Carlson, 1995; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Kutas & Hillyard, 1983).

In the past decade, though, several researchers have argued persuasively for the predictive power of context in generating expectancies during sentence processing. On the view that comprehenders make use of all cues as they become available to constrain upcoming items, with aspects of linguistic input being processed both incrementally and in parallel, preceding context is used to guide the build-up of expectations for particular upcoming items (or item features) based on obligatory syntactic, semantic or phonological dependencies (e.g. Hoeks, Doedens & Stowe, 2000; van Berkum, Kooijman, Brown, Zwitserlood & Hagoort, 2002; Altman & Steedman, 1988; MacDonald, Pearlmutter & Seidenberg 1994; Kamide, Altmann, & Haywood, 2003). Taking into account all the different information sources that may constrain expectancies, there are in turn various levels at which anticipatory processing could occur. For instance, prediction need not be at the level of specific lexical forms, but could instead be at the level of meaning. So although individuals might not be anticipating a specific word, they might pre-activate various semantic features that relate to the word that is most likely to appear in a given context. Alternately (or additionally), linguistic prediction could have a syntactic component, where word classes, thematic dependencies, or grammatical structures are levels at which prediction is occurring.

For instance, Altmann & Kamide (1999) conducted eyetracking experiments using a visual world paradigm. Participants listened to sentences such as "The boy will eat the cake" while viewing a visual scene that contained a ball, a train, a toy car and a cake. As soon as participants heard the word "eat" their eyes moved toward the cake – the only edible object in the scene. Eye fixations were therefore found to focus in on the intended referent well before the word was actually encountered in the sentence. These findings indicated that individuals had used the verb information to select "cake" prior to actually hearing the word. This and other studies (Kamide et al., 2003 and Kamide, Scheepers, Altmann & Crocker, 2002) suggest that syntactic information and semantic constraints together may combine with real-world knowledge to predict what will most plausibly be referred to later in the sentence.

These eye-tracking studies highlight one of requirements for tracking prediction in sentence processing; that is, you need a measure that has high temporal resolution (such as ERPs, magnetoencephalogram or eye movements), and does not alter the comprehension process under study. To this end, some of the recent research into linguistic prediction has used event-related brain potential (ERP) methodology. ERPs reflect the summed extra-cellular potentials of multiple pyramidal neurons in the neocortex acting in unison. Particular advantages of ERPs for studying prediction are that they provide a continuous monitor of brain responses throughout the course a sentence and they provide a measure of the quantitative and qualitative changes that may distinguish two conditions, including sensitivity to semantic and syntactic variables and processing.

2.2.1. The N400 and linguistic prediction

One well-known ERP component, the N400, has been instrumental in developing theories of prediction-based accounts of sentence processing. The N400 ERP component (~200-500 ms post-item onset, more pronounced over posterior scalp sites) was first discovered by Kutas & Hillyard (1980) and has been shown to be the brain's neural response to any potentially meaningful item. The N400's amplitude is especially large to nouns that do not meaningfully fit with their preceding contexts (Kutas & Hillyard, 1980). However, N400s also characterize responses to all but the most highly expected nouns, even when they fit contextually, with amplitudes inversely related ($r \sim -.9$) to their offline cloze probabilities (Kutas & Hillyard, 1984). An item's cloze probability is the percentage of individuals that continue a sentence fragment with that item in an offline sentence completion task.

A few ERP studies in particular have used differential N400 responses to contextually appropriate and inappropriate words to address the question of prediction during language processing. At the level of semantic features, Federmeier & Kutas (1999) have argued strongly for anticipatory processing. They compared the event-related brain potentials (ERPs) to expected sentence completions to those for within and between category violations in highly and weakly constraining sentences in order to determine the extent to which sentence context information is utilized for prediction online. Sentential constraint here was defined by the cloze probabilities of the most expected sentence completions. They found that in high constraint sentences the N400 amplitude to within category violations was reduced relative to low constraint sentences. For example, in the following high constraint sentence pair, the N400 response to *baseball* was reduced relative to that for *monopoly*, such that the response to *baseball* was more similar to that for *football*, the expected exemplar.

(1) *He caught the pass and scored another touchdown. There was nothing he enjoyed more than a good game of...*

football (expected exemplar)

baseball (within category violation)

monopoly (between category violation)

Conversely, in low constraint sentence pairs the N400s to within category violations were not significantly reduced relative to those for the between category violations; in other words, the N400s for within and between category violations were more similar in the low than in the high contextual constraint condition. The N400 amplitudes to expected exemplars and between category violations did not differ significantly as a function of contextual constraint.

What, then, does this imply about prediction? Federmeier & Kutas point out that even though the within category violations in the high constraint sentences had lower cloze probabilities and were rated as less plausible than the within category violations in the low constraint sentences, the N400 in the high constraint condition was smaller. The implication, then, is that N400 amplitude is determined neither by cloze probability nor plausibility alone. Instead, the experimenters suggested that this pattern of results is explained by the fact that the perceptual and semantic features of *baseball* that overlap with *football* were pre-activated in the high constraint condition, because the brain had used the information in the sentence context to access semantic memory and pre-activate (predict) features of the upcoming item.

Van Petten, Coulson, Plante, Rubin & Parks (1999) also used the N400 as a dependent variable for examining prediction during language comprehension. They were interested in investigating whether semantic integration could begin to operate prior to the complete identification of a spoken word. They conducted an auditory ERP study in which participants listened to sentences that were completed by words that were either (a) congruous, (b) incongruous cohorts, or (c) incongruous rhyming words. For instance:

(2) *It was a pleasant surprise to find that the car repair bill was only seventeen...*

dollars (congruous completion)

dolphins (incongruous cohort)

scholars (incongruous rhyming word)

They observed that for *dolphins*, N400 amplitude was low as long as the initial input from the sentence final word was consistent with the sentence context (i.e. for the *dol-* portion of *dollars* and *dolphins* in the example above). However, as soon as the auditory signal for *dollars* and *dolphins* diverged, N400 amplitude to *dolphins* increased rapidly. And importantly, these ERP waveforms' divergence occurred at a time point later than that for the incongruous rhyming word (*scholars*) relative to *dollars*. The authors concluded that semantic processing of a word begins even before it is uniquely identified and that there is a continuous mapping from linguistic input onto semantic representations.

2.2.2. Integration vs. prediction

The results from Van Petten et al. (1999), along with those of Federmeier & Kutas (1999), are in line with a predictive view of sentence processing, and are difficult to reconcile with a purely integrative model of comprehension. However, while both of these studies suggest that language comprehenders pre-activate information about the words that are most likely to be encountered in a given context, the observed effects were evident at the actual target words. These (post-lexical) effects highlight a particular difficulty in studies of anticipatory language comprehension – that is, distinguishing prediction from integration. What some researchers take as evidence for neural pre-activation (prediction, at a psychological level), others take as a sign of the ease or difficulty in integrating words into message level representations upon, but not prior to, their occurrence. As mentioned in the Introduction, N400 amplitude reduction is one finding that highlights this difficulty. For this reason, a clear argument for predictive language procession calls for a design that precludes interpretation in terms of integrative difficulty. A few recent ERP and eye-tracking studies have demonstrated contextually generated expectancies for semantic or syntactic features of upcoming words (Altmann & Kamide, 1999; Tanenhaus, Magnuson, Dahan & Chambers, 2000; Kamide, Scheepers & Altmann, 2003; Chambers & Smyth, 1998; Sedivy, Tanenhaus, Chambers & Carlson, 1999; Wicha, Moreno & Kutas, 2004; Wicha, Bates, Moreno & Kutas, 2003; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005). We will describe some of these in detail below.

2.2.3. Tracking of the time course of prediction

Wicha et al. (2003, 2004) investigated linguistic expectation in Spanish sentences by focusing on the nature and timing of ERPs for nouns of particular gender classes and their preceding gender-marked articles. In separate experiments, native Spanish speakers either listened to or read moderately to highly constraining sentences (completion cloze probabilities ≥ 0.65) that contained a gender-marked article followed by either a written word or an embedded line drawing. The word or line drawing target could be either an expected (high cloze probability) continuation or a semantically incongruent continuation of the same gender class as the expected continuation. In half of the sentences, the gender of the article was incongruent with the gender of the following noun or picture, although participants were not explicitly informed about this. A set of sample Spanish stimuli with their English glosses follow:

Gender Match – Semantically congruous

(3) *Caperucita Roja llevaba la comida para su abuela en una CANASTA muy bonita.*

Little Red Riding Hood carried the food for her grandmother in a[feminine] BASKET[feminine] very pretty.

Gender Match – Semantically incongruous

(4) *Caperucita Roja llevaba la comida para su abuela en una CORONA muy bonita.*

Little Red Riding Hood carried the food for her grandmother in a[feminine] CROWN[feminine] very pretty.

Gender Mismatch – Semantically congruous

(5) **Caperucita Roja llevaba la comida para su abuela en un CANASTA muy bonita.*

Little Red Riding Hood carried the food for her grandmother in a[masculine] BASKET[feminine] very pretty.

Gender Mismatch – Semantically incongruous

(6) **Caperucita Roja llevaba la comida para su abuela en un CORONA muy bonita.*

Little Red Riding Hood carried the food for her grandmother in a[masculine] CROWN[feminine] very pretty.

Even though the article was perfectly appropriate at both a grammatical and semantic level at the point at which it is immediately received, there were ERP effects for gender expectancy at the article, albeit of different polarities in word and picture versions of the experiment: articles preceding words elicited more positivity to violations (which was described as a P600-like response). The P600 is an ERP component that has been associated with the processing of input that represents either a syntactic violation or a disconfirmation of a prior structural choice (Brown & Hagoort, 2000; Hagoort & Brown, 1994). By contrast, the article violations preceding pictures produced a greater negativity relative to congruent articles, an effect more consistent with one in the N400 family, which reflects more semantic processing. In either case, the expectancy effects at the article suggest that individuals were using sentence context to form predictions on-line.

van Berkum et al. (2005) also used gender marking to examine linguistic expectancy for spoken words in Dutch. The test condition in this study involved manipulating the gender marking on prenominal adjectives, rather than articles, as in the Wicha studies. In Dutch, the gender of a noun controls the gender of the inflectional suffix of preceding adjectives. In this study, the experimenters used two-sentence stories that had moderately predictable (greater than 0.50 cloze probability) target noun completions. The sentences were then manipulated so that the inflectional suffix on the adjective preceding the target noun was either congruent with the gender of the highest cloze probability noun that followed, or was of the opposite gender category followed by a lower cloze probability, but still semantically congruent, noun. The following examples are translated from the original Dutch, with two gender classes – neuter and common:

(7) *The burglar had no trouble at all locating the secret family safe. Of course, it was situated behind a ...*(adjective) *but rather unobtrusive...*(noun).

Adjective gender marking consistent with expected noun

a. *een groot...schilderij*

a big[neuter]...*painting*[neuter], no suffix marks neuter gender

Adjective gender marking inconsistent with expected noun

b. *een grote...boekenkast*

a big[common]...*bookcase*[common], -e suffix marks common gender

The low cloze probability noun condition was included to prevent participants from anticipating overt gender agreement violations subsequent to hearing an adjective with unexpected gender marking. For a subset of the sentences with relatively high contextual constraint (completions with a cloze probability of $\geq .75$) van Berkum et al. observed that the adjectives with unexpected gender marking elicited a small but reliable ERP effect relative to those with the expected gender marking. The effect did not begin until the predicted and unpredicted inflectional suffixes of the adjectives began to diverge from each other (e.g. the Dutch word for "big" marked with neuter gender is *groot*, and with common gender *grote*, so the ERP effect began at the point when the pronunciations of these two adjectives begin to diverge). The "unexpected" gender-inflected adjectives (e.g. *grote*[common] when *groot*[neuter] was expected) exhibited a slight positivity between 50-250 ms time-locked to the onset of the adjective inflection. van Berkum et al. attributed the relatively early latency of the effect to the possibility that listeners were sensitive to subtle phonological cues, such as longer vowels, in the stems of the adjectives. They further concluded that the effect at the gender marked adjectives was primarily a syntactic one and that people do indeed use contextual information to form predictions as sentences unfold.

The Wicha and van Berkum studies demonstrate that both readers and listeners generate some linguistic expectations. However, both studies pose some questions that warrant further investigation. Since Wicha et al. were interested in gender mismatches between prenominal articles and nouns, a gender mismatching noun always followed an article of contextually unexpected gender in their design. (Although not all unexpected articles were followed by a noun that mismatched in gender, many were.) Consequently, participants may have picked up on the cue value of the article and thus may have processed it differently than they might have otherwise. By removing all agreement violations it would be more likely that ERPs measured at the article would be indexing a prediction effect related to the expectancy of the article/noun combination rather than, or in addition to, the expectancy of an anticipated agreement violation. van Berkum et al. did essentially this by using context to set up an expectation for a particular noun of a particular gender and violating this expectation at the prenominal adjective, which was then followed by a less expected noun of the same gender. To the extent that they observed an expectancy effect – i.e., when averaging was time-locked to the onset of the inflection (but not the onset of the adjective per se) – it was on an open class word rather than a closed class word as in Wicha et al. It is important to know whether the same results would obtain regardless of word class. Moreover, time-locking to inflectional endings, rather than to the word onsets, makes it difficult to directly compare van Berkum at al.'s results with effects observed in other experiments, such as the Wicha study.

Finally, both the Wicha and van Berkum et al. studies used grammatical gender as a device for investigating prediction in real time language processing, though it would be informative to explore prediction using a linguistic feature other than gender. While grammatical gender marking offers a unique opportunity for exploring expectancy, it is also a linguistically rich phenomenon, operating on syntactic, phonological, and semantic levels simultaneously (indeed, sometimes overtly semantic levels, e.g., feminine gender for biologically female creatures and masculine gender for males.) As such, the very factors that allow it to be a useful device for exploring prediction also make it difficult to isolate its particular role at any given level of language processing. Using a feature other than grammatical gender may therefore provide more precise information about what kind of information, if any, is being pre-activated. Although the English language has few morphosyntactic markings, it does have one characteristic that resembles gender marking in some ways – its use of differing indefinite articles, *a* and *an*, depending on the initial phoneme of the following word. Devoid of gender, case marking and specific semantic content, English indefinite articles offer a means for exploring linguistic prediction at the level of phonological word forms.

2.2.4. The current study

Thus, in pursuit of a clear demonstration for contextual generation of expectancies for specific word forms in semantically meaningful, syntactically well-formed sentences, we designed the current experiment. Our paradigm capitalized on the phonological regularity in English whereby the singular indefinite article meaning "some one thing" is phonologically realized as *an* before vowel sound-initial words and *a* before consonant sound-initial words, e.g., *an airplane* and *a kite*. To determine whether comprehenders pre-activate specific determiners and nouns prior to their occurrence, we used sentences of varying constraint that led to expectations for particular consonant or vowel-initial nouns. Across sentences, target nouns ranged from highly probable to unlikely, based on offline cloze probability norming. For instance, given '*The day was breezy so the boy went outside to fly…*' the most likely continuation was *a kite* (cloze of *a* = 86%, *kite* = 89%). However, the sentence could continue with a plausible, though less likely alternative such as *an airplane*. Based on previous studies we knew that the N400 to *kite* would be smaller than that to *airplane*, and more generally that noun N400 amplitude would be highly inversely correlated with cloze probability. However, as previously mentioned, the pattern of noun effects could be a consequence either of the brain's "surprise" at encountering an item different than what it expects (prediction view), or greater difficulty integrating the received word into the sentence representation (integration view).

Indeed, based on our experiences, *kite* may be easier to integrate into the developing sentence representation than *airplane*. Given the difference in their meanings, it is likely that *kite* and *airplane* also differ in how well they fit with event schemas that the sentence "brings to mind" via semantic memory processes. However, whereas *kite* and *airplane* differ in meaning, *a* and *an* do not, being distinguished only by their phonological forms. Since their semantics are identical and they differ only in frequency of usage and length, there is no reason for the articles to be differentially difficult to integrate into a given sentence representation unless (1) *a* is always easier to integrate, because it is shorter and/or more frequent than *an* in everyday usage, or, as we will maintain (2) comprehenders have already (unconsciously) formed a higher, non-trivial expectation for *kite* than for *airplane*.

2.2.5. Possible outcomes

If anticipation is an integral part of language processing, then it should be reflected in the brain activity probed by the more and less expected indefinite articles. If the amount of pre-activation is driven strictly by word length or frequency, then whatever the ERP effect, it would be context independent, with all *a*'s (versus *an*'s) patterning together. Even if pre-activation is context dependent, the brain may react to the anticipated article with one response and to anything else with a different response, in a binary rather than graded fashion. Finally, if as we hypothesize, consistent with constraint-based models, the language processor exploits sentence context to probabilistically pre-activate possible continuations then this should be reflected in the N400 response to a degree that can be estimated by the eliciting article's offline cloze probability. In sum, no observable difference in the brain's response to more versus less expected articles would be a sharp blow to predictive processing accounts, whereas a graded N400 effect correlated with the article's offline cloze probability would support incremental, predictive processing that is sensitive to meaning-based constraints.

2.3. Methods

2.3.1. Materials

Stimuli consisted of 80 sentence contexts with two possible target types, relatively expected and unexpected indefinite article/noun pairs. Each article/noun pair served as a more and less expected target in different contexts. Targets were sentence medial and congruent (i.e., no agreement violations such as *a airplane*). The 160 stimuli were divided into two lists of 80 sentences, each participant viewing one list. Sentence contexts and article/noun targets were used only once per list. Each list contained equal numbers of relatively expected and unexpected, as well as a and an targets. One quarter of sentences were followed by yes/no comprehension questions. (A full list of the experimental stimuli is included in **Appendix A.**)

In the ERP experiment, different participants read sentences of varying contextual constraint that included target articles and nouns with large ranges of cloze. Across participants the same sentence context appeared with both higher and lower probability articles and nouns. Importantly, although some continuations were more probable than others, none were nonsensical, barring participants from developing a strategy, conscious or unconscious, whereby an improbable article was taken to signal an impending semantic anomaly.

2.3.2. Cloze probability norming

Informed written consent was obtained from all norming (and ERP) participants. Sentence norming was done with different groups of student volunteers, with offline probabilities obtained for all article and noun targets. For articles, cloze ratings were obtained from 30 participants for 80 sentence contexts truncated prior to the target article. For nouns, sentences were truncated following target articles, with two versions of each context (160 sentences total): one with the more probable article supplied, the other with the less probable article. Individual participants saw only one version, with each normed by 30 participants. Participants were asked to provide the best continuations for sentences truncated prior to the article or noun. Article cloze ranged from 0-96%; noun cloze ranged from 0-100%.

2.3.3. ERP participants

Thirty-two volunteers (23 women) participated in the ERP experiment for course credit or for cash. All were right-handed, native English speakers with normal or corrected-to-normal vision, between 18-37 years (mean, 21 years). Seven participants reported a left-handed parent or sibling.

2.3.4. Experimental procedure

Volunteers were tested in a single experimental session conducted in a soundproof, electrically-shielded chamber. They were seated in a comfortable chair approximately one meter in front of a computer monitor and were instructed to read the stimulus sentences for comprehension. They were also informed that some of the sentences would be followed by a yes/no comprehension question, to which they were to respond by pressing one of two hand-held buttons. Response hand was counterbalanced across participants and lists. There was a brief practice session that included sentences with both expected and unexpected targets, as well as filler sentences. Participants were asked to remain still during testing, and to avoid blinking and moving their eyes while the sentences were being presented. Stimuli were presented in 10 blocks of 20 or fewer sentences. The participants were given a short break after each block.

The sentences were presented visually in orange-yellow type on a black background on a cathode-ray tube screen. Each trial began with an empty fixation frame (also orange-yellow in color) appearing in the center of the screen, for a duration that was jittered between 2.5 and 3.5 seconds. The fixation frame stayed on-screen over the course of each sentence, with the sentences presented one word at a time in the center of the frame for a duration of 200 ms with a stimulus onset asynchrony of 500 ms. The fixation frame remained on screen between 1.5 and 2.5 seconds following the offset of the sentence final word, after which a comprehension question either did or did not appear in full on the screen. If a comprehension question did appear, the participant's button-press served to advance the screen to the next sentence. Whether or not there was a comprehension question, there were 2 seconds of blank screen prior to the next sentence automatically appearing on screen.

2.3.5. Electroencephalographic recording parameters

The electroencephalogram (EEG) was recorded from 26 electrodes arranged geodesically in an Electro-cap (**Figure 2.1**), each referenced online to the left mastoid. Blinks and eye movements were monitored from electrodes placed on the outer canthi and under each eye, referenced to the left mastoid. Electrode impedances were kept below 5 KΩ. The EEG was amplified with Grass amplifiers with a band-pass of 0.01 to 100 Hz, continuously digitized at a sampling rate of 250 Hz.

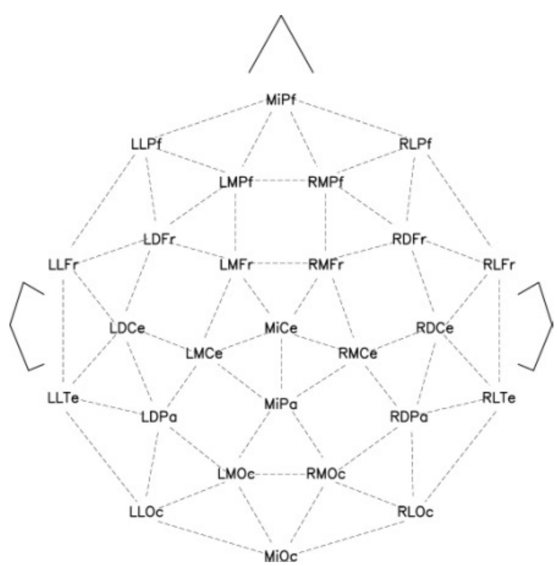


Figure 2.1. Schematic showing the array of the 26 scalp electrodes from which ERPs were recorded.

2.3.6. Data analysis

Trials contaminated by eye movements, excessive muscle activity, or amplifier blocking were rejected offline before averaging – on average, 10.7% of articles and 11.4% of nouns. Data with excessive blinks were corrected using a spatial filter algorithm. A digital band-pass filter set from 0.2 to 15 Hz was used on all data to reduce high frequency noise. Data were re-referenced offline to the algebraic sum of left and right mastoids and averaged for each experimental condition, time-locked to the target article and noun onsets.

2.4. Results

2.4.1. Behavioral Results

Comprehension accuracy was calculated for the yes/no probe questions. Participants correctly answered an average of 94% (range = 88% to 100%) of the questions, indicating that they were attending to and comprehending the experimental sentences during the recording session.

2.4.2. ERP Results

ERPs were analyzed for target articles and nouns in 160 sentences, with 16 participants viewing each item. Motivated by previous findings of N400 amplitude inversely correlated with cloze probability (Kutas & Hillyard, 1984), we performed two main types of analyses.

First, we conducted more traditional ANOVA analyses across all electrode sites on the mean amplitudes of the articles and noun targets classified as 'high' (≥50% cloze) or 'low' (<50% cloze) as determined in the cloze norming. Any significant interactions with electrode site were reported only when of theoretical significance, and in those instances we followed up with distributional analyses using a subset of 16 representative channels across the scalp. These locations were divided into: 2 levels of Hemisphere (left vs. right) X 2 levels of

Laterality (lateral vs. medial) X 4 levels of Anteriority (prefrontal vs. frontal vs. parietal vs. occipital). See **Figure 2.2** for electrode locations. Where appropriate, Huynh-Feldt (HF) epsilon correction was performed, and reported throughout are the original degrees of freedom and the HF-corrected *p*-values.

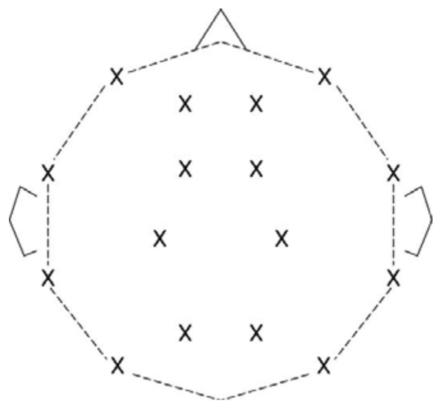


Figure 2.2. Locations of the 16 electrode sites used for the distributional analyses.

Second, the broad ranges of article and noun expectancy allowed us to conduct correlation analyses, for more fine-grained examination of the relationship of the ERP effects and the offline probabilities of the relevant continuations. For our correlations, the 160 articles and nouns were sorted into ten equal-width bins as a function of each item's cloze probability, from highest (90-100%) to lowest (0-10%). ERPs for each 10% bin were averaged first within, then across, participants. The average numerical cloze probability of each bin was then calculated and correlated with mean ERP amplitude in the N400 time window (200-500 ms), separately for articles and nouns. Correlation coefficients (*r*-values) and percentage of variance explained by offline probability (*r*-squared) were then calculated separately for all 26 electrode sites. Following the recommendations of Nakagawa (2004), we report for this experiment – and all subsequent studies included in this thesis – statistically significant correlations deemed so according to their exact (i.e., uncorrected for multiple comparisons) *p*-values. Our intent in doing so is to allow for the reader's evaluation of the biological importance (and statistical significance) of these results. Statistically significant ($p \leq .05$) correlation values are indicated with an asterisk (*) on the figures with *r*-values plotted topographically on scalp maps.

In addition to analyses of the N400 time window for both the target articles and nouns, visual inspection suggested that in a time window following (or perhaps overlapping with) the noun N400, at least at some electrode sites, there was an increased positivity to low relative to high cloze nouns. For this reason, we also measured noun mean amplitudes over a later time window (500–1200 ms).

2.4.2.1. N400 ERP effects

2.4.2.1.1. Nouns

Performing a binary sort on noun cloze probability, we conducted an ANOVA with 2 levels of Cloze (high vs. low) X 26 levels of Electrode. As expected, N400 amplitude decreased (became less negative) as noun cloze probabilities increased (**Figure 2.3**).

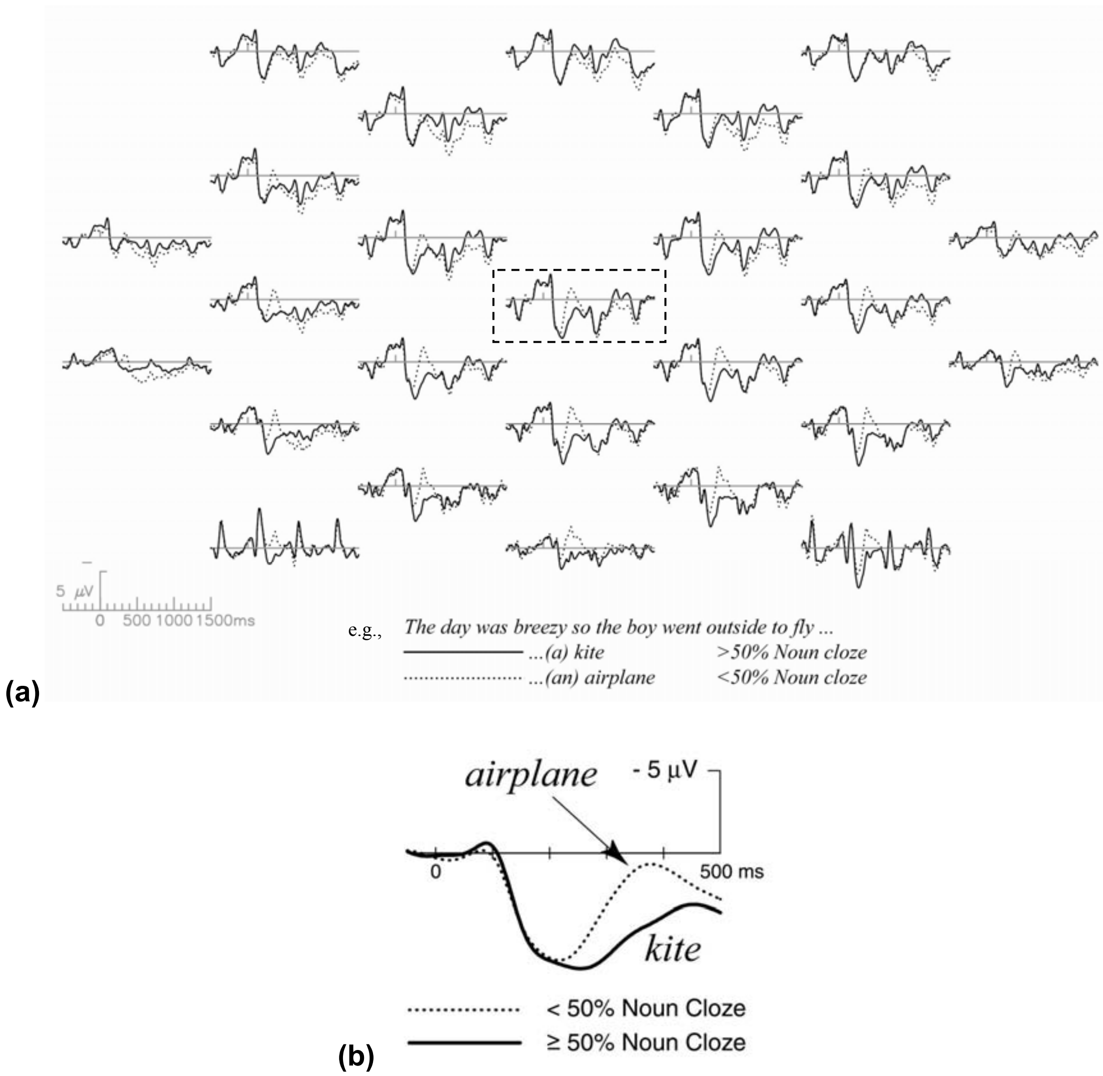


Figure 2.3. Grand average nouns sorted on noun cloze probability. Solid lines represent nouns with >50% cloze, dotted lines represent nouns with <50% cloze. Negative amplitudes on this and all subsequent ERP figures are plotted upwards. The noun N400 cloze effect plotted over (a) all 26 channels, with vertex electrode highlighted, and (b) at the vertex electrode.

There was a main effect of Cloze, $F(1,31) = 22.93$, $p < .0001$, with mean amplitude of low cloze nouns (.56μV) relatively more negative than that of high cloze nouns (1.98μV). There was also an interaction of Cloze X Electrode, $F(25,775) = 25.65$, $p < .0001$, which was explored further by conducting a distributional analysis. Distributional analyses (as described above) were conducted via an omnibus ANOVA with the following factors: 2 levels of noun Cloze X 2 levels of Hemisphere (left vs. right) X 2 levels of Laterality (lateral vs. medial) X 4 levels of Anteriority (prefrontal vs. frontal vs. parietal vs. occipital). The Cloze effect had the general distributional pattern characteristic of the N400: largest over medial and posterior sites, with a right-lateralized skew. While there were significant interactions of Cloze with all other factors (Hemisphere, Laterality and Anteriority), these were mediated by a significant higher-level interaction of all four factors (*F*(3,93) = 3.52, *pHF* = .0355). This interaction revealed a pattern of similar N400 effects at both left and right medial sites, increasing in amplitude from frontal to more posterior sites. This anteriority progression was also present at left and right lateral sites, but with smaller effect sizes (smallest at left lateral sites). In fact, at left lateral sites there was even a reversal in the N400 effect at the prefrontal and frontal sites, with low cloze nouns being more positive than high cloze items (see Figure 2.4).

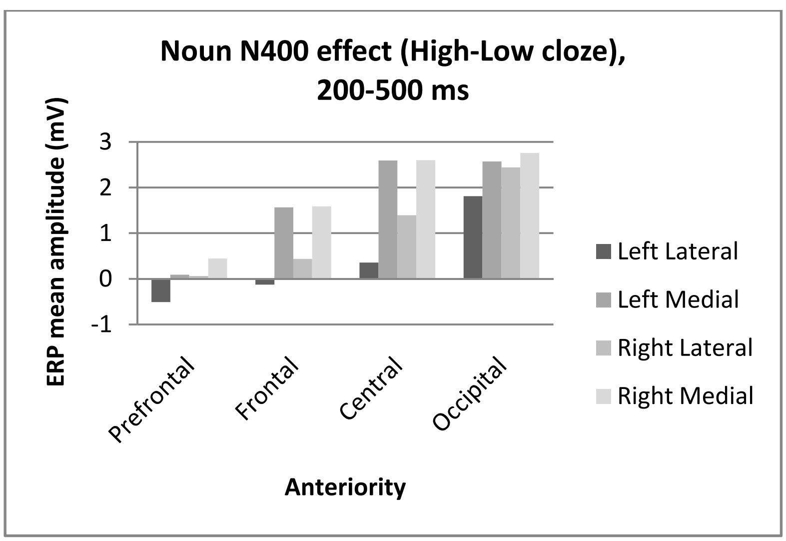


Figure 2.4. Distributional analysis of noun cloze probability effect (>50% cloze minus <50% cloze), N400 time window: Interaction of Cloze X Hemisphere (2 levels) X Laterality (2 levels) X Anteriority (4 levels).

In addition to the main effect of cloze probability and subsequent distributional interactions, we replicated the well-known correlation between N400 amplitude and offline cloze of the target nouns (**Figure 2.5a**), with correlation coefficients ranging from $r = 0.36$, n.s., to $r = -0.84$, $p < 0.01$ at various scalp sites (**Figure 2.5b**). Noun cloze probability thus accounted for up to 71% (r^2) of variance in brain activity between 200-500 ms after a noun's appearance. Moreover, correlations peaked over posterior sites where N400 amplitudes are typically largest, whereas anterior sites (where visual N400s are usually less prevalent) exhibited little if any evidence of similarly correlated brain activity (**Figure 2.5b**). These results were an important precondition for analysis of the articles because they demonstrated that the different degrees of constraint in these materials were reflected in offline expectancies and N400 amplitude modulations in the usual way. However, as previously noted, the noun correlation pattern does not settle the question of prediction because high correlations could reflect either the degree of pre-activation or the variance in the integrability of the noun with the mental representation of the sentence up to that moment.

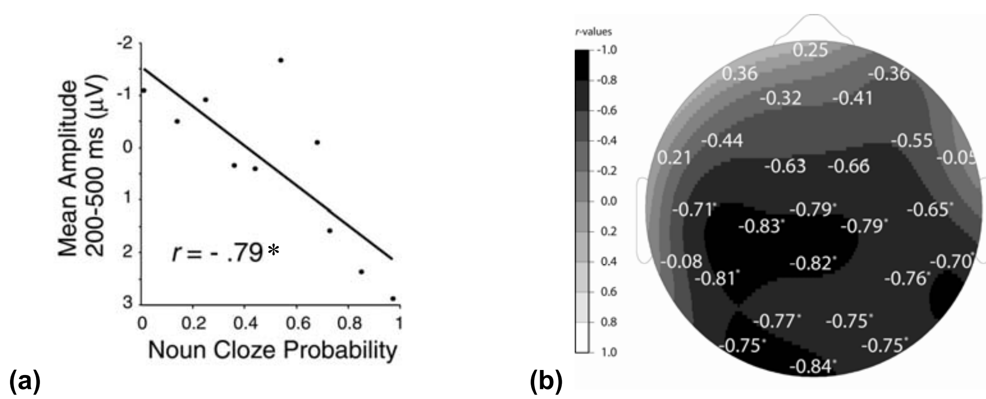


Figure 2.5. Mean noun N400 amplitude correlated with noun cloze probability. (a) At the vertex electrode, the scatter diagram shows a strong inverse relation between cloze and N400 amplitude, *r* = -0.79, *p* < 0.01. The best fitting regression line is also plotted. (b) The *r*-values for all 26 electrode sites are plotted on an idealized head, looking down, nose at the top. Darker shading indicates larger negative correlations, with *r*-values between sites estimated by spherical spline interpolation. The dotted circle demarcates the vertex. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

2.4.2.1.2. Articles

To address the question of prediction directly we conducted both ANOVA and correlation analyses in the article N400 time window (200-500 ms). Although article ERP waveforms were significantly smaller than those elicited by nouns, visual inspection suggested that the amplitude of the negativity in the article N400 time window might vary as a function of article expectancy (**Figure 2.6**). The ANOVA with 2 levels of article Cloze probability (high vs. low) X 26 levels of Electrode revealed mean amplitude of low cloze articles (-0.18μV) to be slightly more negative than high cloze articles (0.08μV), however this difference was not significant [$F(1,31) = 0.93$, $p = 0.34$], and there was no interaction of article Cloze with electrode site [$F(25,775) = 0.97$, $p = 0.43$].

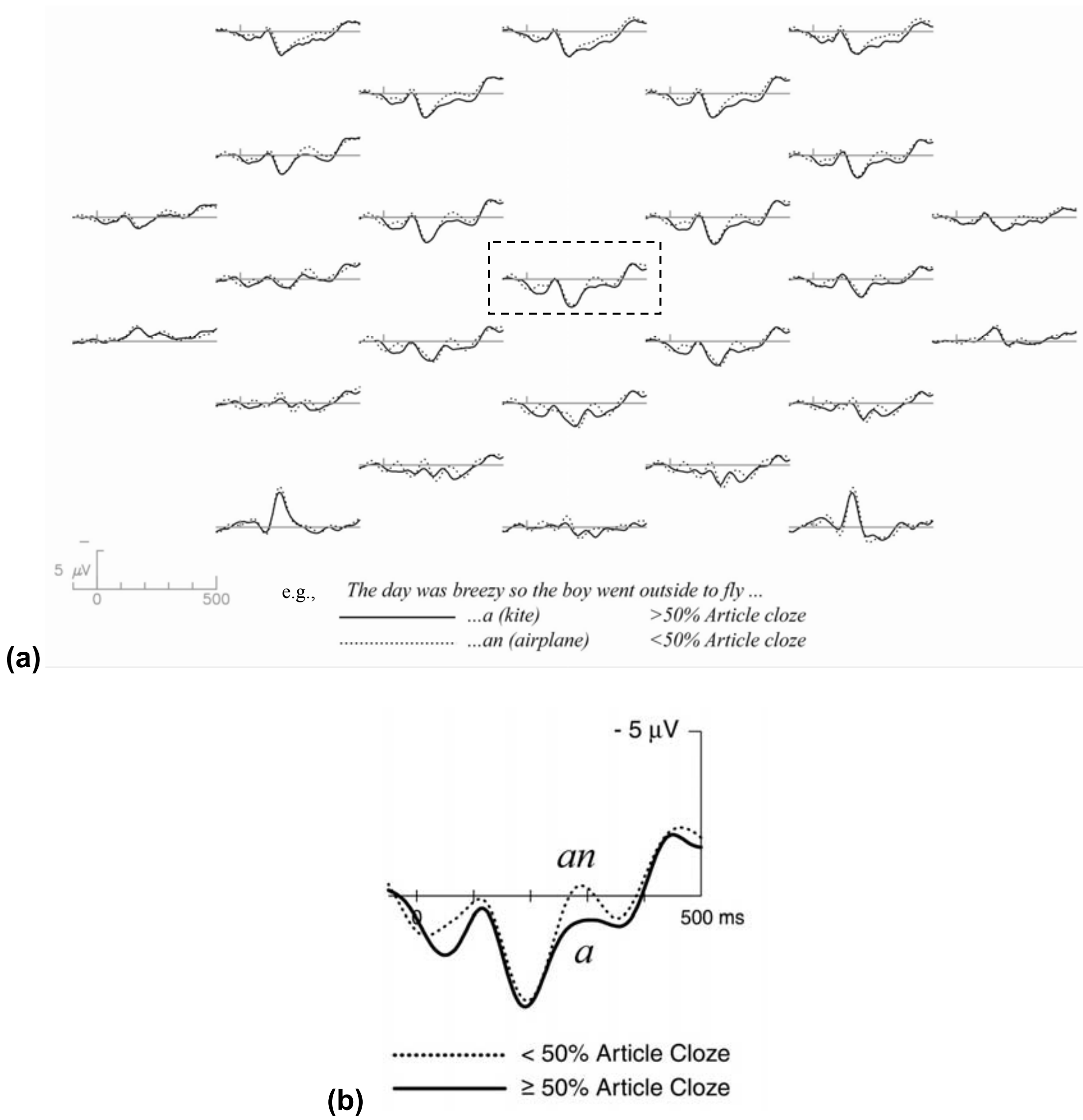


Figure 2.6. Grand average articles sorted on article cloze probability. Solid lines represent articles with >50% cloze, dotted lines represent articles with <50% cloze. The article N400 cloze effect plotted over (a) all 26 channels, with the vertex electrode highlighted, and (b) at the vertex electrode.

To determine whether article mean amplitude and article cloze probability patterned together, we turned to the correlation analysis, using the ten 10% article cloze bins. Just as for the nouns, the higher the articles' cloze probability, the smaller the ERP negativity between 200-500 ms post-onset (Figure 2.7a), with correlation coefficients ranging from r = -0.19, n.s., to r = -0.72, p < 0.05, at various recording sites (Figure 2.7b).

Moreover, maximum correlations clustered over centro-parietal sites, similar to nouns, albeit somewhat more right lateralized (contrast **Figures 2.5b** and **2.7b** and see following section). So at least over certain scalp areas, up to 52% of variance in article N400 amplitude was accounted for by the average probability that individuals would continue the sentence context with that particular article offline.

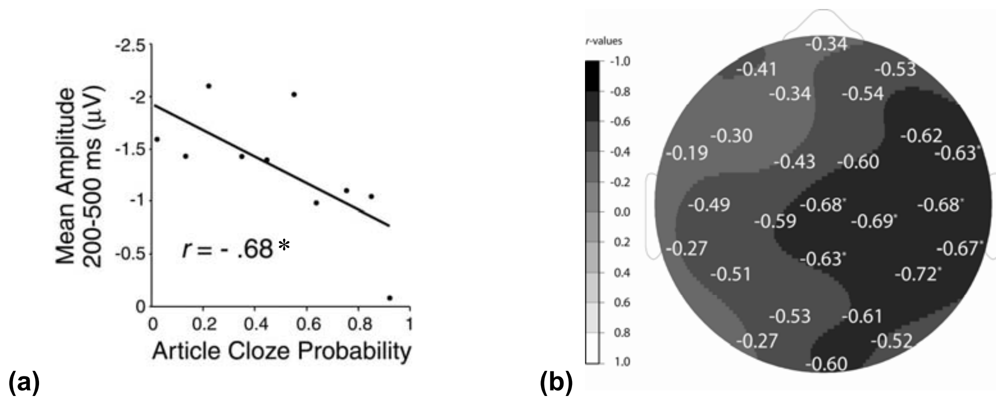


Figure 2.7. Mean article N400 amplitude correlated with article cloze probability. (a) At the vertex electrode, the scatter diagram shows a strong inverse relation between cloze and N400 amplitude, *r* = -0.68, *p* < 0.05. The best fitting regression line is also plotted. (b) The *r*-values for all 26 electrode sites are plotted on an idealized head, looking down, nose at the top. Darker shading indicates larger negative correlations, with *r*-values between sites estimated by spherical spline interpolation. The dotted circle demarcates the vertex. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

2.4.2.1.3. Contrasting noun and article N400 effects

To explore and contrast the spatial distributions of the noun and article N400 effects (200-500 ms post onset), a three-way repeated measures ANOVA with Subjects as a random factor and the factors Word Type (articles, nouns), Cerebral Hemisphere (left, right) and Anteriority (anterior, posterior) was conducted using 5 electrode sites per scalp quadrant (20 locations total, see **Figure 2.8**).

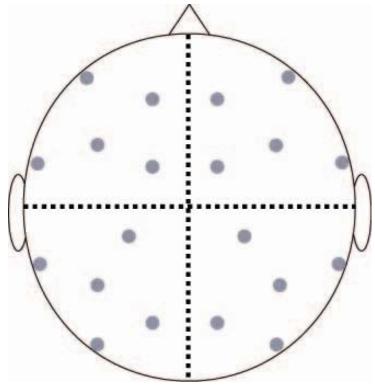


Figure 2.8. 20 electrodes used for comparisons of spatial distributions of ERP effects.

The results revealed a Word Type effect ($F_{1,31} = 6.67$, $p < 0.05$) with prediction effects larger for nouns than articles, an overall Hemisphere effect ($F_{1,31} = 14.58$, $p < 0.001$) with prediction effects larger over right than left hemisphere scalp sites, and an overall Anteriority effect ($F_{1,31} = 25.06$, $p < 0.0001$) with larger effects observed over posterior than anterior scalp sites. There was also a two-way interaction between Word Type and Anteriority ($F_{1,31} = 15.98$, $p < 0.001$), with noun effects significantly larger than those for articles over the posterior but not anterior sites.

An analysis of the scalp distribution of the 20 corresponding correlation measures in **Figures 2.5b** and **2.7b** was also conducted using the same three factors as above, after applying Fisher's *r*-to-*Z* transformation to express the correlation values in units normalized by their standard deviation (*Z*-scores). The ANOVA did not reveal a main effect of Word Type ($F_{1,8} = 0.62$, $p = 0.45$, n.s.), but there were overall effects for Hemisphere ($F_{1,8} = 20.66$, $p < 0.01$) and Anteriority ($F_{1,8} = 6.97$, $p < 0.05$), with larger correlations over the right than left hemisphere, and larger correlations over posterior than anterior sites. Post-hoc Bonferroni-corrected pairwise comparisons of the three-way interaction (Word Type, Hemisphere, Anteriority) revealed that for the articles the correlations showed a significant right hemisphere bias over both anterior ($p = 0.0004$) and posterior ($p = 0.0008$) sites. For the nouns, however, this hemispheric pattern was present only at anterior sites (*p* = 0.0015), and was not reliable over posterior sites (*p* = 0.11), where the correlations were more symmetrical.

2.4.2.2. Noun late positivity (LP)

2.4.2.2.1. High vs. low cloze probability nouns

The results of the distributional ANOVA within the noun 200-500 ms time window indicated that the typical N400 pattern (larger negativity to low cloze relative to high cloze nouns) "flipped" over a subset of electrode sites – primarily at left, lateral, frontal scalp locations. Visual inspection also indicated that low cloze nouns had a more positive mean amplitude than high cloze nouns in a prolonged time window beyond the N400. To investigate this observed pattern, mean amplitudes were measured for the high and low cloze conditions and ERP effects were assessed by means of repeated measures ANOVA across all electrode sites. (Note: For analyses performed over the late time window(s), 'cloze effect' will refer to the increased positivity to low relative to high cloze nouns.)

Because observation of the waveforms suggested the cloze effect was quite extended, analyses were first conducted over a 500-1200 ms time window. Between 500-1200 ms, low cloze nouns were found to be more positive than high cloze nouns [$F(25,775) = 5.74$, $p = .0228$], with Cloze interacting with Electrode site [$F(25,775) = 7.00$, $pHF = .0002$]. Distributional analyses revealed an interaction of Cloze X Anteriority that was mediated by Laterality [$F(3,93) = 21.19$, $pHF = <.0001$], in which low cloze nouns were more positive than high cloze nouns at central, frontal, and prefrontal sites, with progressively larger effects toward the front of the head at medial sites (maximal cloze effect at prefrontal site was -1.80 μV). For lateral sites, Cloze effects were largest at central scalp locations (-.83 μV).

We also wanted to determine if the cloze effect between 500-1200 ms might be reflecting a greater ERP contribution of a shorter time window within the 500-1200 ms one.

To test this, we performed similar mean amplitude analyses using 500-800 ms (early LP) and 800-1200 ms (late LP) time windows. Main effects of Cloze (greater positivity to low cloze nouns) were observed in both time windows, [$F(1,31) = 4.26$, $p = .0474$] and [$F(1,31) = 5.74$, $p = .0228$] respectively, and for both there were interactions of Cloze X Electrode site, which were followed up with distributional analyses.

For the 500-800 ms time window there was an interaction of Cloze X Hemisphere, mediated by Laterality [$F(1,31) = 7.66$, $p = .0094$] indicating that the largest cloze effects were present at left lateral sites (**Figure 2.9**). There was also an interaction of Cloze X Anteriority that was mediated by Laterality [$F(3,93) = 35.03$, $p = <.0001$] with an increasingly large cloze effect from the back to front of the head over medial sites (maximal cloze effect was -2.00 μV at prefrontal sites), but no such progression over lateral channels (the maximal cloze effect was -1.11 μV at central sites, smaller at prefrontal, frontal, and occipital sites), see **Figure 2.10**.

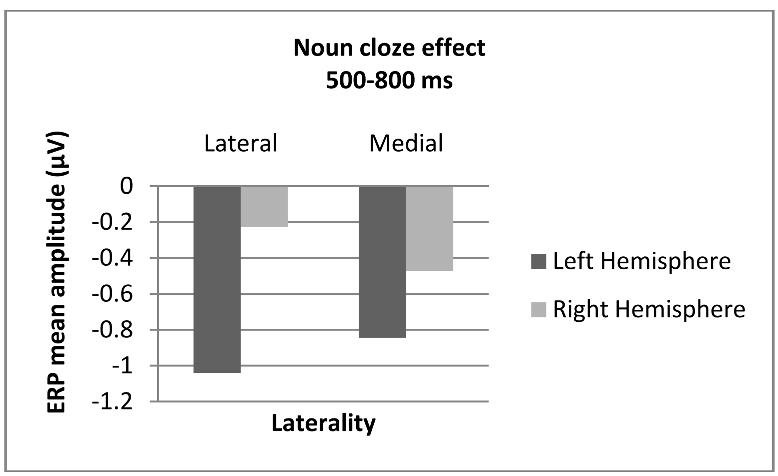


Figure 2.9. Distributional analysis of the noun cloze probability effect (>50% cloze minus <50% cloze), 500-800 ms time window: Interaction of Cloze X Hemisphere (2 levels) X Laterality (2 levels).

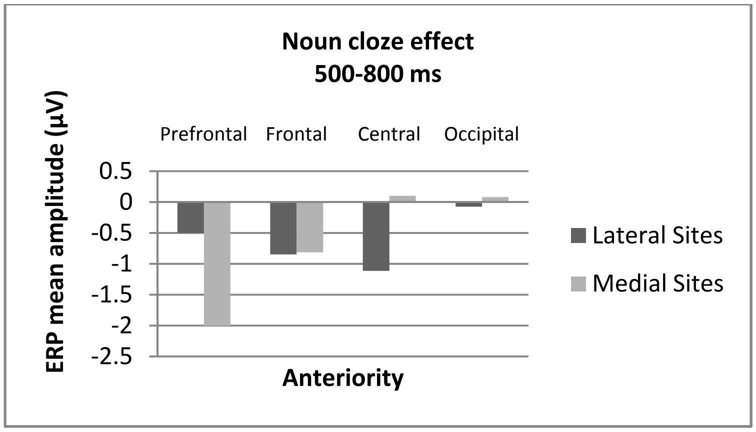


Figure 2.10. Distributional analysis of noun cloze probability effect (>50% cloze minus <50% cloze), 500-800 ms time window: Interaction of Cloze X Laterality (2 levels) X Anteriority (4 levels).

For the 800-1200 ms distributional analysis the main effect of Cloze was only marginally significant [$F(1,31) = 4.01$, $p = .0540$], but there were significant interactions of Cloze X Laterality X Anteriority, which were mediated by a higher order interaction of the three factors (Cloze X Laterality X Anteriority), [$F(3,93) = 10.01$, $pHF = 0.0002$]. Similar to the 500-800 ms time window, Cloze effects increased progressively from back to front at medial sites (maximal effect was -1.66 µV at prefrontal sites), while effects at lateral sites were maximal at central electrodes (-.63 µV), see Figure 2.11.

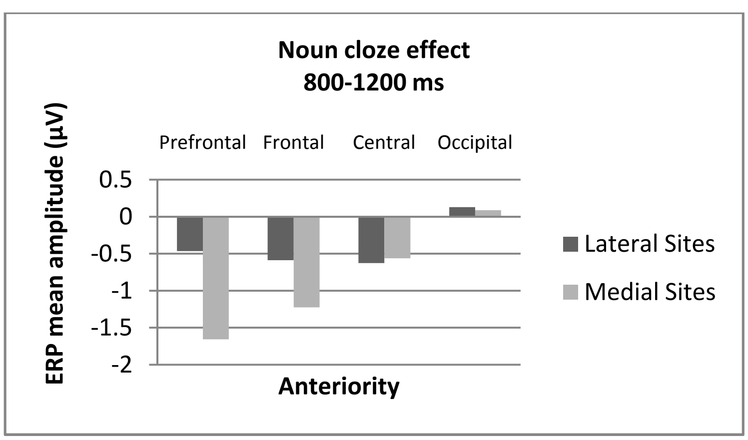


Figure 2.11. Distributional analysis of noun cloze probability effect (>50% cloze minus <50% cloze), 800-1200 ms time window: Interaction of Cloze X Laterality (2 levels) X Anteriority (4 levels).

To compare the spatial distribution of the LP cloze effects in the early (500-800 ms) and late (800-1200 ms) LP time windows, a repeated measures ANOVA with Subjects as a random factor and the factors Time Window (early, late), Hemisphere (left, right) and Anteriority (anterior, posterior) was conducted using 5 electrode sites per scalp quadrant (20 locations total), see **Figure 2.8**. There was no main effect of Time Window, but there was a main effect of Anteriority [$F(1,31) = 9.25$, $p = .0048$] with larger cloze effects at frontal sites (-1.05 µV) than at posterior sites (-.30 µV). There was also an interaction of Time Window with Hemisphere [$F(1,31) = 20.44$, $p = .0001$] indicating a more left lateralized cloze effect in the 500-800 ms time window than between 800-1200 ms (**Figure 2.12**).

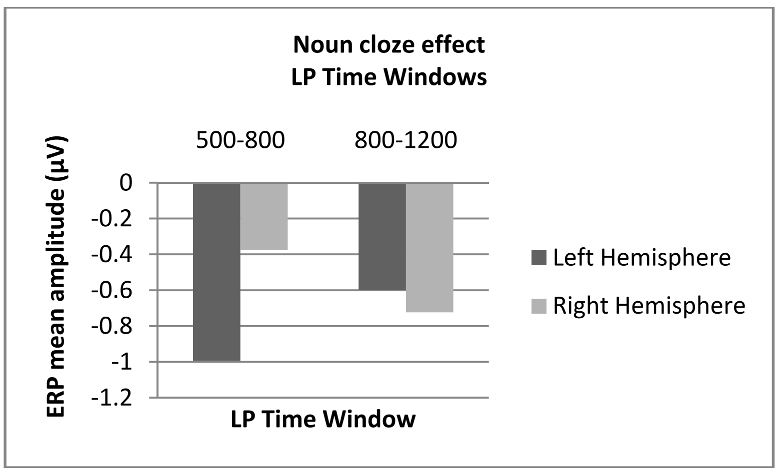


Figure 2.12. Distributional analysis of noun cloze probability effect (>50% cloze minus <50% cloze), LP time windows: Interaction of Cloze X Hemisphere (2 levels) X Time Window (2 levels).

2.4.2.2.2. Correlations of noun LP with noun cloze probability

Similar to the analyses done in the N400 time window, we also ran correlations on noun cloze probability and noun ERP mean amplitude within the three late positive time windows (**Figure 2.13a**, **b**, and **c**). Though none of the correlation values reached statistical significance, in all three time windows (500-1200, 500-800, 800-1200 ms) the patterns of correlations tended to be more negative, increasing from posterior to anterior scalp sites.

Unlike the N400 time window then, as cloze probability decreased, ERP mean amplitude became more positive. A comparison of the scalp distribution of the 20 corresponding correlation measures (**Figure 2.8**) in the early and late LP time windows was also conducted using the same three factors as above, after applying Fisher's *r*-to-*Z* transformation to express the correlation values in units normalized by their standard deviation (*Z*-scores). The ANOVA revealed a main effect of time window [$F1,1 = 9.84$, $p = 0.0037$], with the later LP time window having larger negative correlations. (We speculate that this difference could, in part, be due to residual ERP activity in the N400 time window overlapping and therefore "dampening" the correlation strength in the early LP time window.) There was also an overall effect of Anteriority ($F1,1 = 8.79$, p =0.0057), with larger negative correlations over the anterior than posterior sites.

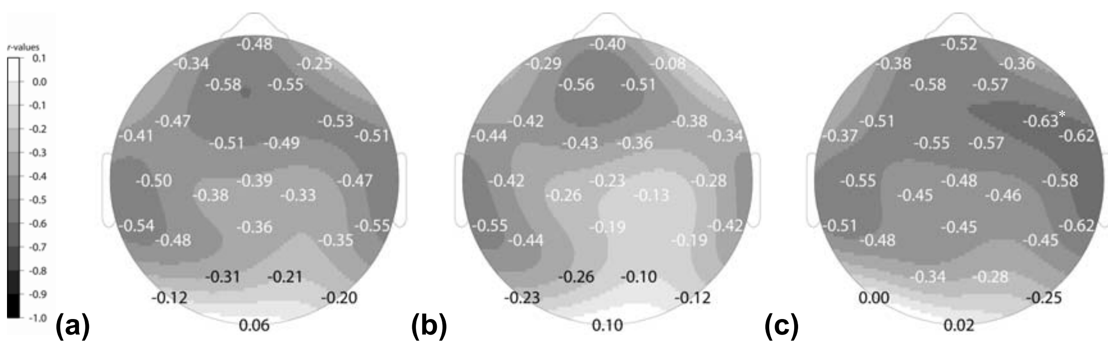


Figure 2.13. Correlations of noun cloze probability and ERP mean amplitude in three LP time windows: (a) 500-1200 ms, (b) 500-800 ms, and (c) 800-1200 ms. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

2.4.2.2.3. Summary of noun LP results

In sum, analyses of the LP time windows revealed a clear effect of cloze probability – low cloze nouns more positive than high cloze nouns – between 500-1200 ms post noun onset. This effect increased in amplitude from the back to front of the head over medial sites and was largest at central electrodes over lateral sites: this pattern held within both of subtime windows, 500-800 ms and 800-1200 ms. Additionally, there was a left hemisphere bias between 500-800 ms, with the most pronounced effects observed over lateral sites. Correlation analyses also revealed a tendency for increasing ERP positivity as noun cloze decreased, more so over anterior sites, with larger (though non-significant) negative correlations in the later than earlier LP time window.

2.5. Discussion

2.5.1. N400 and prediction effects

By constructing sentence contexts that led to varying offline expectations for vowel or consonant sound-initial nouns, we could assess the extent to which such expectations were formed online by preceding the noun with the phonologically appropriate indefinite article or the other semantically identical, sententially congruent, but phonologically inappropriate one. Similar to the nouns, the more contextually unexpected an indefinite article was, the more negative the ERP mean amplitude between 200-500 ms post-word onset (N400). In other words, the brain's response to the articles differed in a graded fashion as a function of contextual constraint. Our results thus demonstrate not only that readers can rapidly, incrementally integrate incoming words into evolving mental sentence representations, but that they do so, in part, by exploiting various constraining forces to form probabilistic predictions of which specific words will come next. Here, we clearly showed this for the target articles and nouns, though we have no reason to assume the same would not hold for every word in a sentence throughout the range of normal reading rates. Notably, maximum correlations for both nouns and articles were not randomly distributed across the scalp but rather clustered over centro-parietal scalp sites (**Figures 2.5b** and **2.7b**) where previous reading studies have revealed the largest N400 effects. This topographical pattern indicated that the values were not simply the spurious outcome of multiple testing at the 26 electrode sites.

In conducting the correlational analyses, we established the functional relationship of the negativity between 200-500 ms post article onset to the canonical N400 typically elicited by nouns and verbs. Critically, we also demonstrated that this negativity indexed expectancy for the eliciting article (and upcoming noun). Given that all articles were grammatically and semantically congruent within their contexts and that *a* and *an* have identical semantics, there was no reason for either article type to have been any more or less difficult to integrate into the sentence representation. Systematic variation in amplitude of the ERP negativity in relation to offline article cloze probability thus constitutes strong evidence that participants were indeed anticipating the phonological form of a particular noun and therefore had formed expectations for one article type relative to the other, and apparently experienced some processing difficulty when the less expected article appeared.

Articles are relatively short, highly frequent, highly predictable as a word class, not as semantically rich as nouns, verbs, adjectives, or adverbs, and are often skipped over during natural reading (O'Regan, 1979). In addition to providing unequivocal evidence for lexically specific prediction then, the article correlations are compelling, perhaps surprising, evidence that articles too are predicted and integrated with context in qualitatively similar ways as nouns. For reasons not yet known, the correlations with offline probability were on average lower for articles than nouns (although at some electrode sites, the two were statistically similar). Nonetheless, the article correlations clearly demonstrated that prediction is not limited to highly constraining contexts. We believe that this sort of anticipation is an integral (perhaps inevitable) part of real time language processing, and is likely to play a functional role, though this has yet to be demonstrated.

Our findings thus suggest that individuals can utilize linguistic input that becomes available to them to pre-activate representations of upcoming words in advance of their appearance. Exactly what informs these predictions, as well as the neural mechanism for predictive language processing, are matters for empirical and computational investigations. An open question, for instance, is how the human sentence comprehension system handles variation in natural input rates (e.g., 2-3.5 words/sec), and in particular, whether the same (or different) mechanisms are engaged. In line with most studies of comprehension that draw general conclusions without systematically varying input rate, we assume that basic language processing mechanisms do not vary fundamentally across the range of normal input rates. This parsimonious assumption is bolstered by results of N400 studies in which variation in presentation rate revealed no evidence for the engagement of qualitatively different neural mechanisms (Kutas, 1987; Gunter, Jackson & Mulder, 1992). While the current study only demonstrated graded prediction at the slower end of natural input rates, we suggest that this conclusion may generalize to faster rates, given the aforementioned arguments and ERP evidence for binary prediction (expected versus unexpected) in natural speech (Wicha, Bates, Moreno & Kutas, 2003; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005). Subsequent experimentation will undoubtedly shed more light on these issues.

We propose that words in a sentence, alone and in combination with others, operate via semantic memory to tap into and thereby differentially activate information, going beyond the immediate physical input. Semantic memory is presumed to include information about individual words as well as world knowledge built up from experience. It is this experiential knowledge about people, places, things, and events accessed by the linguistic input from which we maintain that probabilistic pre-activation of particular word forms follows. Our observation of an ERP expectancy effect at the article leads us to conclude that predictions can be for specific phonological forms – words beginning with either vowels or consonants. In this sense, we propose that prediction can at least under some circumstances be highly specific.

Our results are in line with a growing list of empirical studies demonstrating that the brain's language parser projects probabilistic expectations about various aspects of linguistic processing during online sentence comprehension tasks. Findings from several of these studies have been taken as evidence that the parser uses constraints accruing as a sentence is analyzed word by word to: 1) compute likely relationships among referents in linguistic and visual contexts (e.g., upon hearing the word *eat*, a person is likely to scan the environment for something edible), (Altmann & Kamide, 1999), 2) pre-activate semantic features of categories (e.g., expecting a particular kind of tree pre-activates features of trees, even when not all trees would be plausible in the sentence context), (Federmeier & Kutas, 1999), or 3) anticipate various syntactic aspects of to-be-presented material (e.g., expecting the grammatical gender of upcoming items in gender-marked languages such as Spanish or Dutch) during word-by-word reading and natural speech (Wicha, Moreno & Kutas; 2004; Wicha, Bates, Moreno & Kutas, 2003; van Berkum, Brown, Zwitserlood, Kooijman, Hagoort, 2005).

In particular, our study expands in several critical and novel ways upon findings from the aforementioned grammatical gender studies, by Wicha and colleagues and van Berkum and colleagues, where nouns were preceded by words whose syntactic gender marking was inconsistent with that of the "expected" noun. Although both the gender studies and our study used the same general logic, the differences between the experimental manipulations, designs, and analyses lead to substantive differences in the conclusions that can justifiably be drawn. Whereas the Spanish and Dutch studies utilized pre-nominal gender marking on determiners and adjectives, respectively, our study relied upon a purely phonological (sound representation) relation between probed articles and upcoming nouns. Thus, our observation of article ERP variation provides a strong test of whether the language system predicts word forms with specific phonological content (lexemes), instead of simply representations specifying words' semantic and syntactic properties (lemmas). In addition, having tested prediction with semantically identical *a/an* articles (function words) instead of words richer in meaning (content words), such as adjectives, we effectively counter the argument that the observed difference between more versus less predictable articles reflects difficulty interpreting them. And perhaps most importantly, only our study compared brain activity elicited by a range of more or less predictable articles, not simply most versus least expected. The article correlation findings thus show for the first time that the language system does not simply pre-activate a single word when its representation exceeds some threshold given a highly constraining context. Instead, a gradient of pre-activation shows that the system makes graded predictions.

2.5.2. Late positivity (LP) effects

In the current study, designed to test for lexical prediction by examining ERPs to more and less expected prenominal indefinite articles (*a*/*an*), there was a surprising finding in the form of a prolonged, late frontal positivity (500-1200 ms) to unexpected relative to expected nouns (e.g., *airplane* in '*The day was breezy so the boy went outside to fly an airplane…*'). Though correlations of noun cloze probability and LP mean amplitude indicated a pattern of negative *r*-values – i.e., increasing LP with decreasing cloze – the factor of cloze probability does not seem to fully account for this ERP effect. This is evident when the weaker correlation values in the LP time window are contrasted to the strong (and differentially distributed) correlations in the noun N400 time window. So although cloze probability may be indirectly related to the LP effect, it seems unlikely that the two components are reflecting (different stages of) the same functional process.

What to make then of the LP following the N400 to contextually less expected nouns? Certainly, as a fortuitous finding, any interpretation of these results is purely speculative, and could be best informed by other similar ERP findings. Complicating the matter is the fact that late positivities accompanying N400s to semantic experimental manipulations have been only sporadically reported in the language ERP literature, most frequently to incongruent sentence completions. However, there has been little (if any) systematic meta-analysis of the variety of conditions under which post-N400 positivities obtain. (For instance, LPs have been found by Coulson & Van Petten, 2002, to metaphoric sentence endings; by Swick, Kutas & Knight, 1998, to incongruent sentence completions; by Moreno, Federmeier & Kutas, 2002, to English-Spanish code switches and low cloze sentence continuations; by Coulson & Wu, 2005, to unrelated probe words following control sentences compared to the same probes related to joke sentences.) Additionally, the more frontal distribution of the LP observed in our study marks its contrast to a variety of other LPs which have been extensively discussed in the sentence processing literature. In particular, more posterior effects include the P600 (aka Syntactic Positive Shift or SPS) observed in response to syntactic violations and ambiguities, and the more recent proliferation of studies noting P600s to so-called "semantic violations" (for a review, see Kuperberg, 2007). These "semantic P600s" have been observed to such manipulations as animacy violations (Kuperberg, Sitnikova, Caplan, & Holcomb, 2003), semantic verb argument violations (Kim & Osterhout, 2005), and semantic reversals (Kolk, Chwilla, van Herten & Oor, 2003), to name a few. Not to mention that in our study, there were no "violations" per se; rather, we used perfectly congruent endings which were simply contextually less expected.

The LP to unexpected nouns detectable as early as the N400 time window and extending to 1200 ms is a novel finding in the present study. Because our results at the article constitute strong evidence for contextual preactivation of specific word forms, we believe that the LP effect may also be related to anticipation and what happens in the face of receiving linguistic input that does not correspond with a highly predicted item. As an initial hypothesis for the source of this effect, we would like to offer the following proposal.

When reading sentences such as those in our study, an unexpected article might not immediately discount the possibility of still receiving the expected noun, since the noun could potentially be preceded by an adjective whose initial phoneme is consistent with the article presented. So if expectations were violated at the article, participants may nonetheless choose (unconsciously) to maintain their original noun prediction by means of an adjective. In this case, when the unexpected (low cloze probability) noun eventually appears, not only would it be semantically more difficult to integrate into the context (hence the N400), but it would also be surprising at a syntactic (word class) level if the language processor were anticipating an adjective in order to be able to salvage its original noun prediction. This account of the frontal positivity at the noun would be in line with P600 effects that have been noted for the increased processing costs at points of syntactic ambiguity or structural unexpectedness (e.g. Osterhout & Holcomb, 1992 and Brown & Hagoort, 2000). To illustrate our point, here is an example stimulus item from our study.

(8) *Dale was sorry for what he had said to Bernadette and knew that he owed her a...because she was upset.*

...a check... (unexpected)

....an apology... (expected)

One can imagine that when presented with the unexpected article (in this case *a*), it may be difficult for the parser to immediately activate an alternative continuation, given that *apology* had been so strongly preactivated. Instead, the parser may activate a phoneme-consistent intermediary adjective to maintain its original noun expectation (e.g., *sincere* or *heartfelt apology*) when there is not another noun continuation at the ready. Once an adjective is preactivated, comprehenders would then be "surprised" by the appearance of *check* (or any other noun) in that position. This outcome would be reflected in a posterior N400 reflecting the difficulty of integrating the word into the sentence representation, possibly coincident with a frontal positivity reflecting the mismatch with expected word class. This account offers a testable hypothesis for the late positive ERP effect at the noun following unexpected articles. A first step to following up on this proposal would be to determine in a systematic way the kind of responses comprehenders supplied offline, in the sentence norming task already conducted, when asked to complete sentences that continue with unexpected indefinite articles (see Experiment 3A). Until additional testing of the LP can be performed we must withhold further speculation.

2.6. Summary and conclusions

Our electrophysiological results extend previous prediction findings in several important ways. First, they demonstrate that a candidate entity (or its depiction) need not be physically present in order for the brain to narrow the possibilities for likely continuations; rather, predictions can emerge based upon associations that form as sentential context accrues. Second, our results illustrate that at least one subclass of function words (which generally provide more grammatical structure than lexical meaning) – indefinite articles – can play an important role in building context and facilitating linguistic processing. This finding is particularly relevant given the paucity of evidence in the comprehension literature for semantic context effects on function words (King & Kutas, 1995; Kluender & Kutas, 1993). Third, our LP finding suggests that there may be some consequence to strongly preactivating but not receiving a particular word or token of a particular word class. A possible future direction and extension of the present research, then, would be to test offline "recovery" strategies and determine how they potentially map onto the LP ERP effect. And finally, our findings unambiguously show that anticipatory processing can be not only for conceptual or semantic features, but for specific phonological word forms. In sum, although natural language comprehension must occur over a range of input rates, with a nearly infinite number of possible word combinations, these factors do not appear to prevent the brain from anticipating the most probable continuations for sentences.

2.7. Appendix A. Experimental stimuli used in Experiment 1

Eighty experimental sentence contexts were used in the ERP study, each continuing with one of two possible indefinite article (*a* or *an*) + noun pairs, for a total of 160 unique sentence stimuli. For each context, one of the two target continuations was more likely than the other, though there was a range of cloze probabilities for the more and less likely continuations across contexts. As listed below, the first continuation is the more probable one, while the second continuation is less likely. Article + noun pairs that serve as probable continuations in one context serve as less probable ones in a "paired" context. The complete set of stimuli, with sentence pairs in contiguous order, is listed below:

1) The old wives' tale says that if you want to keep the doctor away then you should eat <u>an apple/a carrot</u> a day.
2) For the snowman's eyes the children used two pieces of coal, and for its nose they used <u>a carrot/an apple</u> from the fridge.
3) It was difficult to understand the foreign professor because he had <u>an accent/a lisp</u> when he spoke.
4) Katie did not like to say words with the letter "s" because she spoke with <u>a lisp/an accent</u> and was embarrassed.
5) Dale was very sorry for what he had said to Bernadette and he knew that he owed her <u>an apology/a check</u> because she was still upset.

6) The bakery did not accept credit cards so Peter would have to write a check/an apology to the owner.

7) Although the idea of flight was as old as the hills, the Wright Brothers were the first people to build an airplane/a kite that actually flew.

8) The day was breezy so the boys went outside to fly a kite/an airplane in the park.

9) Surrounded by mountains, Lola shouted across the valley and heard an echo/a bird in the distance.

10) The tweeting in the treetops sounded like a bird/an echo to Melissa.

11) When the scuba diver saw the tentacle, he quickly realized that the creature under the rock was an octopus/a fish in hiding.

12) Marge hated baiting the hook, but she knew that it was the only way to catch a fish/an octopus without using a lure.

13) On a street corner in Paris Kim had her portrait painted by an artist/a student for three francs.

14) Betsy had no desire to enter the working world, so decided to stay at the university and remain a student/an artist for a few more semesters.

15) Her grandfather always stressed how important it was for Sophia to go to school, because he had never had the opportunity to receive an education/a book while growing up.

16) James is an avid reader, so for his birthday his sister decided to give him a book/an education about Africa.

17) When the representative retired in the middle of his term, the state was forced to hold an election/a parade in his district.

18) Without the floats and the marching bands it wasn't much of a parade/an election this year.

19) When her husband started staying late at the office every night, Joanne began to suspect that he was having an affair/ a fight with someone at work.

20) Whenever Josh had too much to drink, he became belligerent and would try to start a fight/an affair with someone in the bar.

21) Instead of giving her money whenever she asked for it, Rachel's father thought that it would be a real lesson in economics for her if every week she received an allowance/a dollar to spend as she wished.

22) Andrew complained that the only place he could afford was the 99-cent store, after his grandmother gave him a dollar/an allowance for helping her.

23) When she was filled out the paperwork, Merrill had to write down the name of someone who could be contacted in case of an emergency/a fire while she was at work.

24) Harris saw smoke billowing out of his neighbors' home and immediately got on the phone to report a fire/an emergency at their address.

25) Charlie's wife told him that she was tired of pressing his shirts and that he would have to learn how to use an iron/a hanger one of these days.

26) I wanted to put my coat in the closet but I could not find a hanger/an iron anywhere.

27) No matter how safely you drive your car, chances are someday you'll be involved in an accident/a race with another vehicle.

28) When we saw the finish line and the pylons sectioning off part of the road, we knew that there must have been a race/an accident earlier that day.

29) Margot had finally decided to buy a house in the suburbs, after a year of renting an apartment/a studio in the city.

30) Instead of recording a live album, the band decided they would have more control if they recorded in a studio/an apartment without a lot of background noise.

31) Sue had wanted to go to Tim's birthday party but she was still waiting for an invitation/a message from him.

32) When I called his house Nolan was not home, but his mother said she could relay a message/an invitation to him.

33) The highlight of Jack's trip to India was when he got to ride an elephant/a bicycle in the parade.

34) You never forget how to ride a bike/an elephant once you've learned.

35) Because it frequently rains in London, it's a good idea to always carry an umbrella/a newspaper with you.

36) As he walked past the corner drugstore on his way to work, Carl skimmed the headlines and decided to go ahead and buy a newspaper/an umbrella when he saw what the forecast was.

37) Frank wanted to design a very modern house so he sought advice from an architect/a plumber to select the fixtures.

38) When the pipe broke in the bathroom, Felicia looked through the phonebook to find a plumber/an architect who could come take a look at the job.

39) Because Bart did not clean his wound properly, he ended up getting an infection/a tattoo on his leg.

40) The Hell's Angel rolled up the sleeve of his leather jacket to show the girls where he had gotten a tattoo/an infection the previous night.

41) In order to get into an R-rated movie, children under 17 must be accompanied by an adult/a parent or legal guardian.

42) Older children often have a harder time than younger children dealing with the loss of a parent/an adult who has cared for them.

43) Jeffrey mailed the letter without a stamp/an envelope, so the post office would not deliver it.

44) When Wendy went to pay for the birthday card, the clerk pointed out that she had forgotten to pick up an envelope/a stamp along with it.

45) The professional photographer was disqualified from the photo contest because the rules explicitly state that you must be an amateur/a beginner in order to compete.

46) There are always ways to improve your game, whether you're an advanced tennis player or merely a beginner/an amateur starting out.

47) Amelia did not want to go to the bar with her friends after seeing the movie, so she made up an excuse/a story for why she needed to go home early.

48) Before little Gloria went to sleep, she wanted to hear a story/an excuse from her mother.

49) At first Victoria did not know why her brother was crying over the sink, but then she noticed that he had just sliced an onion/a finger with his knife.

50) Marie wanted to sample a tiny bit of the sauce so she daintily dipped a finger/an onion into the pot.

51) After Joanne's first book was published, she finally felt like she could call herself an author/a success in front of her peers.

52) Alicia's first client was a failure, but her second was a success/an author and made her a lot of money.

53) Lance had moved to Hollywood in hopes of becoming an actor/a director, but so far he had only done a few commercials.

54) From a young age Steven had an interest in filmmaking and had always dreamed of being a director/an actor when he grew up.

55) Dana enjoyed singing in private, but was scared to death at the thought of performing in front of an audience/a crowd at the concert.

56) Carrie's father had always joked that two is company but three is a crowd/an audience for him.

57) Damon preferred climbing stairs over riding in an elevator/a wheelchair because it gave him a chance get a little exercise.

58) The doctors would not allow Monica to walk so soon after her surgery, so the only way for her to move from place to place was in a wheelchair/an elevator accompanied by a nurse.

59) The pilot had to make an emergency landing in the middle of the desert because he was nowhere near an airport/a city or a safe place to land.

60) Hannah wanted to live in a small town, but her husband preferred to live closer to a city/an airport because of his job.

61) Although the basketball team's defense was very strong, they did not have much of an offense/a coach since the middle of the season.

62) Ever since Mr. Barnes had moved away, Bobby's football team had been left without a coach/an offense and thus lost all their games.

63) The group had been brainstorming all day but they still didn't have an idea/a name for their project.

64) The guys didn't know what to call their band, so Trent told everyone to come up with a name/an idea and they would vote.

65) Being from the mainland, Karen never got used to the feeling of living on an island/a ship and being so removed from everything.

66) Sandra decided that she would work for a cruise line for a year before starting college, although she had never been on a ship/an island and did not know how to swim.

67) As Elizabeth climbed the ladder in the barn, she heard a hooting sound coming from the rafters and looked up to see an owl/a cat staring down at her.

68) Every time they went for walks, Sylvia's dog Rex would break into a run as soon as he spotted a cat/an owl, but luckily Rex couldn't climb trees.

69) The first time Timmy saw the Pacific, he thought it was a lake, but his brother laughed and explained to him that it was an ocean/a mirage and he should be careful of the waves.

70) As the men stumbled across the desert, they thought they saw an oasis on the horizon, but it turned out to be a mirage/an ocean so they continued walking.

71) The waitress at the next table was trying to scribble everything down, but it was obvious that she did not know how to take an order/a compliment from her customers.

72) Marcy did not deal well with praise and had never really learned how to take an order/a compliment from her superiors.

73) Carly wasn't sure if the Venus flytrap was classified as a plant or as an animal/a flower, but it seemed to have characteristics of both.

74) While Natasha was strolling through the colorful gardens, she reached down and picked a flower/an animal up off the ground.

75) The chemistry teacher explained to the class that water and salt were both compounds but that nitrogen was just an element/a liquid that hadn't been combined.

76) Orlando thought that glass was a solid but when he held it under the Bunsen burner it melted and turned into a liquid/an element that could be sculpted.

77) Violet was just learning how to make an omelet, so her father began by showing her how to crack an egg/a window in the kitchen.

78) Because they were playing baseball so close to the house, the children ended up shattering a window/an egg and destroying a nest that some birds had built on the sill.

79) Bob claims to be from outer space, but nobody believes that he is actually an alien/a tourist from another planet.

80) With his khaki shorts, his loud Hawaiian shirt, and the camera around his neck, you could definitely tell that the man was a tourist/an alien in the country.

2.8. References

Altmann, G.T.M. & Steedman, M.J. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

Altmann, G.T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Bloom, P.A. & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, 8, 631-642.

Boland, J.E., Tanenhaus, M.K., Garnsey, S.M. & Carlson, G.N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34, 774-806.

Brown, C. & Hagoort, P. (2000). On the electrophysiology of language comprehension: Implications for the human language system. In M. W. Crocker, M. Pickering & C. Clifton Jr. (Eds.), *Architectures and mechanisms for language processing.* (213-237). New York, NY, US: Cambridge University Press.

Carroll, J. & Bever, T. (1978). The perception of relations: a case study in the relation of knowledge and perception. In: Carterette, E., Friedman, M. (Eds.), *Handbook of perception, Vol. 7, Speech and language*. Academic Press, New York.

Chambers, C.G. & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language*, 39, 593-608.

Coulson, S. & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, 30(6), 958-968.

Coulson, S. & Wu, Y.C. (2005). Right hemisphere activation of joke-related information: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17(3), 494-506.

Daneman, M. & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 9, 561-584.

Federmeier, K.D. & Kutas, M. (1999). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Gunter, T.C., Jackson, J.L. & Mulder, G. (1992). An electrophysiological study of semantic processing in young and middle-aged academics. *Psychophysiology*, 29(1), 38-54.

Hagoort, P. & Brown, C.M. (1994). Brain responses to lexical ambiguity resolution and parsing. In: Clifton, Jr., L. Frazier and K. Rayner, Editors, *Perspectives On Sentence Processing*, Erlbaum, Hillsdale, NJ, USA, 45-80.

Hoeks, J., Doedens, G. & Stowe, L. (2000). Lexical Prediction and Semantic Relatedness Interact in Sentence Processing: Evidence From ERPs. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, 30 March - 1 April 2000.

Kamide, Y., Altmann, G.T.M. & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133-156.

Kamide, Y., Scheepers, C. & Altmann, G.T.M. (2003). Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, Vol. 32, No. 1, 37-55.

Kamide, Y., Scheepers, C., Altmann, G.T.M. & Crocker, M.W. (2002). Integration of syntactic and semantic information in predictive processing: Anticipatory eye-movements in German. Presented at the 14th Annual CUNY Conference on Human Sentence Processing, New York, USA, in March 2002.

Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205-225.

King, J. W. & Kutas, M. (1995). A brain potential whose latency indexes the length and frequency of words. *CRL Newsletter* (University of California, San Diego) 10, 1-9.

Kintsch, W. & van Dijk, T.A. (1978). Towards a model of text comprehension and productions. *Psychological Review*, 85, 363-394.

Kluender, R. & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8, 573-633.

Kolk, H.H.J., Chwilla, D.J., van Herten, M. & Oor, P.J.W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1-36.

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

Kuperberg, G.R., Sitnikova, T., Caplan, D. & Holcomb, P.J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17(1), 117-129.

Kutas, M. & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539-550.

Kutas, M. (1987). Event-related brain potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. In *Current Trends in Event-Related Brain Potential Research, EE*G Suppl. 40 (eds. Johnson, R., Rohrbaugh, J.W. & Parasuraman, R.) 406-411 (Elsevier, Amsterdam).

Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161-163.

MacDonald, M.S., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*. 101(4), 676-703.

Marslen-Wilson, W.D. & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

Marslen-Wilson, W.D. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226-228.

Mitchell, D.C. & Green, D.W. (1978). The Effects of Context and Content on Immediate Processing in Reading. *Quarterly Journal of Experimental Psychology*, 30, 609-636.

Moreno, E.M., Federmeier, K.D. & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, 80(2), 188-207.

O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics*. 25(6), 501-509.

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

Pickering, M.J. (1994). Processing local and unbounded dependencies: A unified account. *Journal of Psycholinguistic Research*, 23, 323-352.

Schwanenflugel, P.J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instrumentation, and Computers*, 18, 363-371.

Sedivy, J.C., Tanenhaus, M.K., Chambers, C.G. & Carlson, G.N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-148.

Stanovich, K.E. & West, R.F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition*, 7, 77-85.

Steedman, M. (1989). Grammar, interpretation, and processing from the lexicon. In: Marslen-Wilson, W. (Ed.), *Lexical Representation and Process*. MIT Press, Cambridge, MA.

Swick, D., Kutas, M. & Knight, R.T. (1998). Prefrontal lesions eliminate the LPC but do not affect the N400 during sentence reading. *Journal of Cognitive Neuroscience Supplement*, 29.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D. & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, 29(6), 557-580.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632-1634.

Traxler, M., Bybee, M. & Pickering, M. (1997). Influence of connectives on language comprehension: eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology A*., Vol 50A(3), 481-497.

Tyler, L.K. & Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 683-692.

van Berkum, J.J.A., Brown, C. M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

van Berkum, J.J.A., Kooijman, V., Brown, C.M., Zwitserlood, P. & Hagoort, P. (2002). Do listeners use discourse-level information to predict upcoming words in an unfolding sentence? An ERP study. Annual meeting of the Cognitive Neuroscience Society (CNS-2002), San Francisco, April 14-16, 2002. *Supplement to Journal of Cognitive Neuroscience* (p. 82).

Van Petten, C., Coulson, S., Plante, E., Rubin, S. & Parks, M. (1999). Timecourse of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 25 (2): 394-417.

Wicha, N.Y.Y. (2002). Dissertation: Grammatical Gender in Real-Time Language Comprehension in Spanish: Behavioral and Electrophysiological Investigations.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

A significant portion of Chapter 2 is comprised of material from an article in *Nature Neuroscience*, Volume 8(8), Aug 2005, 1117-1121, DeLong, Katherine A., Urbach, Thomas P. & Kutas, Marta, for which I was the primary author and investigator.

CHAPTER 3.
EXPERIMENT 2: TESTING THE LIMITS OF PREDICTION: EFFECTS OF PRESENTATION RATE ON LINGUISTIC PRE-ACTIVATION

3.1. Abstract

In Experiment 1 of this dissertation, an ERP prediction effect – in the form of a cloze probability-graded negativity in the N400 time window – was observed at prenominal indefinite articles preceding more and less expected sentence continuations. Because N400 amplitude decreases are thought to reflect the degree to which an item has been semantically facilitated by a particular context, this finding was interpreted as evidence that comprehenders had already formed expectations for upcoming nouns and experienced difficulty when they encountered prediction-inconsistent articles. One potential concern with these findings was that our visual presentation rate (2 words per second) may have allowed participants "spare time" to predict relative to the average natural rate of reading (closer to 4 words per second on some accounts, e.g., Rayner, 1998). Since the brain may have a limited capacity to pre-activate information based on the timing of the input, the prediction argument would be strengthened by demonstrating anticipatory effects at faster input rates. The current study thus investigated whether there was evidence for predictive language processing at a visual input rate of 3.3 words/sec, using a design otherwise identical to Experiment 1. Our results from the current study showed, similar to Experiment 1, N400-like prediction effects at the prenominal articles, though in the present study the N400-like effect was graded across only a subset of participants and was associated with the cloze probability of the subsequent target noun, rather than the article itself. Also similar to Experiment 1, we observed a late ERP positivity (LP) to unexpected nouns (following the typical N400 effect for such words) - an effect that we suggest may also relate to predictive language processing. We conclude by speculating on possible explanations, some theoretical

and others methodological, for the differences and similarities between the current study and Experiment 1.

3.2. Introduction

The question of whether readers and listeners actively predict upcoming words or whether they simply interpret words with respect to the preceding context after the words are presented is one that has been explored through various methodologies over the past several decades. The question has become something of a hot topic lately, likely garnering renewed interest for a number of reasons. First, these two processing views are ultimately grounded in very different concepts of how the language comprehension system works, and with technological developments offering a variety of new ways for probing brain functions, the field may now be better equipped to investigate the neural underpinnings of meaning construction. Along with these methodological advances, researchers may also be better poised to integrate computational models of prediction with data from human subjects, including evidence from corpora analyses. Additionally, there seems to be a convergence across neuro-scientific subfields to view the brain as a sort of “prediction machine”, with this view extending across both levels of neural processing as well as functional domains. Finally, it may be that anticipation (aka prediction, preactivation, and anticipation) is a phenomenon that most comprehenders feel like they already believe in – it seems intuitive. So for psycholinguists, the zeitgeist within the larger neuroscience community may provide an additional impetus to challenge long-standing skepticism of anticipatory models based on Chomskian theories of generativity.

Tracing the background of anticipatory language processing research, comprehension studies have shown that semantic context effects are ubiquitous across a number of dependent measures (e.g., accuracy, speed, electrical brain activity) and a variety of experimental tasks. Across behavioral measures, the general finding has been that items preceded by supportive semantic contexts are processed more quickly and accurately than those not preceded by such contexts, and when perceptually or semantically ambiguous items are encountered, comprehenders tend to select context-consistent alternatives. Other evidence comes from event related brain potential (ERP) studies, where the N400 component consistently exhibits a gradation in amplitude inversely correlated with the offline likelihood of encountering an item in a particular context (i.e., the cloze probability derived from norming data). On a "prediction" view facilitation of probable items reflects the benefit of prior information, and on an "integration" view slower reaction times (RTs) in behavioral tasks and larger N400s when reading for comprehension both reflect interpretative difficulty when an item does not "fit" as well with an already-formed semantic representation of the context. However, as simple as the prediction/integration question may seem, it has been a challenging one to investigate, given that it is very difficult to "capture" prediction as it is happening, rather than inferring it after the fact. Thus, while there has been evidence of various sorts that language comprehenders might have been predicting during language processing, there is very little direct empirical evidence that they did in fact do so.

ERPs (as well as eye movements) have proven especially useful in addressing this issue because they afford a look at brain (or eye movement) activity as language input is being processed – before, during and after context is found to have a beneficial or a detrimental effect. For instance, Wicha and colleagues (Wicha, Moreno & Kutas, 2003a, 2004; Wicha, Bates, Moreno & Kutas, 2003b) and van Berkum and colleagues (e.g., Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Otten, Nieuwland & van Berkum, 2007) have both used the ERP methodology with essentially the same logic as we did in Experiment 1 of this dissertation: first, set up contextual constraints such that a particular noun is expected, and then violate this "expectation" not (or not just) at the noun but rather at some point before the noun. The pre-nominal expectation violation is required because any effect at the noun could just as easily be attributed to differential integration of the word with its context upon its occurrence as to pre-activation of the noun prior to its occurrence. While both the Wicha and van Berkum studies capitalized on a grammatical gender system (in Spanish and Dutch, respectively) to probe for prediction by preceding the noun with a word whose syntactic gender marking was inconsistent with that of the syntactic gender of the "expected" noun, in Experiment 1 we manipulated the phonological expectancy of upcoming nouns by way of English indefinite articles (*a/an*). Given that at the prenominal probe points in the sentences, there is nothing wrong at any representational level (with the exception of a gender or phonological mismatch with a possible continuation), a differential effect in the ERP could be taken as evidence for some sort of prediction. All three research groups observed differential prenominal ERP effects that were offered as evidence of prediction, though the Wicha and van Berkum findings were based on binary conditions (more versus less expected), compared to cloze probability-graded stimuli and graded ERP effects at the prenominal article in our Experiment 1. Notably, some of the grammatical gender studies used auditorily presented natural speech stimuli – a method that many would argue is more naturalistic than the rapid serial visual presentation (RSVP) used in our study.

A lingering question generated by our visually-presented Experiment 1 then – and an issue often cited as a concern for language comprehension ERP research in general, where sentences are frequently presented at RSVP rates slower than those normally encountered during natural language processing – is the issue of mapping results to "online" language processing. For instance, in our study, sentences were presented at rate of two words per second. Although in ERP studies, methodological considerations are usually the motivation for using these slower presentation rates (to allow for minimal overlap and thus more straightforward interpretation of the brainwaves associated with a particular event), normal reading rates are often considered to be closer to 3-4 words per second (Rayner, 1998). Thus, a potential argument against slower rates of presentation is that they may allow participants "spare time" to predict, with some perhaps questioning the generalizability of such findings to more natural language processing.

While it is indisputable that sentence comprehension can indeed proceed at rates faster than the presentation rates that we used, by some estimates our 2 words/sec presentation rate is not so slow. This rate compares favorably with the slow end of average rates according to the 130-190 words/min that Reynolds and Givens (2001) use to approximate the average speaking range and the 189-231 words/min Lewandowski, Codding, Kleinmann & Tucker (2003) estimate as the reading rate for college students (i.e., an average of 2 to 3.5 words/sec). So although some might reserve the label of "real time" processing for referring to faster rates, comprehension occurs effortlessly and routinely over a range of input rates.

There are, however, important general questions about how human sentence comprehension mechanisms handle variation in input rates and, in particular, whether the same processes are engaged over the entire range of rates. As far as we know, it is an open question whether input rate variability is accommodated by modulating a single set of sentence comprehension mechanisms or by engaging different systems under different conditions. The methodologically parsimonious working hypothesis adopted by many psycholinguists is that the basic mechanisms do not vary fundamentally as a function of input rate – an assumption that would seem to be shared (perhaps tacitly) by other sentence comprehension studies that seek to draw general conclusions without systematically varying presentation rates. In ERP studies that have varied presentation rate, there has been no ERP evidence for the engagement of different neural mechanisms (e.g., Kutas, 1987; Gunter, Jackson, & Mulder, 1992).

With regard to the study of predictive language comprehension, though, the issue of timing seems particularly relevant, as there exists the possibility that prediction of lexical forms does not generalize to faster input rates. Against this speculation, though, stands both our empirical evidence from Experiment 1 that prediction does indeed occur at least at the slower end of normal input rates, along with methodological and empirical support for the working hypothesis that sentence processing mechanisms do not fundamentally differ at different input rates (within reason). Methodological differences notwithstanding, the auditory (natural speech) data from the Wicha et al. (2003) and van Berkum et al. (2005) studies also show binary expectancy (expected versus unexpected) effects in the ERP, suggesting that timing is not an important determinant of whether or not effects are obtained.

Nonetheless, one direct way – the way that we will pursue here – of testing the possibility of whether graded effects of linguistic prediction obtain at faster input rates is to use the same experimental design and stimulus materials as Experiment 1, except using an increased rate of presentation, i.e., decreased stimulus onset asynchrony, or SOA. (Another way, similar to the Wicha and van Berkum studies, would be to use natural speech stimuli, presented in the auditory modality.)

3.2.1. Possible results

To review, Experiment 1 demonstrated the following: a prediction effect at the pre-target article (e.g., '*The day was breezy so the boy went outside to fly an airplane in the park*'), observable as an N400-like negativity whose amplitude was inversely correlated with the expectancy of the articles, and a pronounced N400 effect at the noun, with mean amplitude correlating with the off-line cloze probability of the nouns. Unexpected nouns also exhibited a late ERP positivity (from 500-1200 ms) relative to more expected nouns – a finding that we suggested might relate to some type of expectancy violation. Although previous research (e.g., Kutas, 1987) indicates that N400 onset and peak latency to content words (nouns) are similar regardless of presentation rates (with the exception of words presented at an extremely rapid rate of 10 per second), it is an empirical question whether predictive N400-like effects will be observed to function words such as articles when the presentation rate approaches that of natural language. If pre-activation of particular phonological word forms holds at sentence presentation rates closer to that of natural language processing, then we might expect to observe ERP effects at both the articles and nouns similar to our findings in Experiment 1. These results would argue against the idea that the article prediction effects in Experiment 1 were a consequence of "extra" processing time. If however, the brain has a limited capacity to pre-activate certain kinds of information based on the timing of the input, then we might expect not to observe any ERP effects at the article.

In addition to examining article prediction effects at the faster presentation rate, we also intend to monitor ERP effects at the noun, both in the N400 as well as in the late positivity (LP) time windows. We anticipate that the noun N400 will continue to exhibit a graded sensitivity to the cloze probability of the more and less expected nouns. And though n400 effects following presentation of the target noun make difficult the task of distinguishing between prediction and integration views, in our analyses and discussions in Study 1, we have proposed that the LP effect may reflect some aspect of failed prediction. Observing similar increases in LP amplitude to low cloze noun continuations in the current study would indicate that whatever cognitive function the LP may be indexing, this processing related to comprehension is consistent across various input rates.

3.3. Methods

3.3.1. Materials and procedure

The materials (see **Appendix A**, Experiment 1) and procedure used for this experiment were identical to those used for Experiment 1, except that the RSVP sentences were presented at a rate of 1 word every 300 ms, instead of every 500 ms. Individual words were visible on screen for 200 ms with a 100-ms inter-stimulus interval (ISI), yielding a stimulus onset asynchrony (SOA) of 300 ms. See **Figure 3.1**.

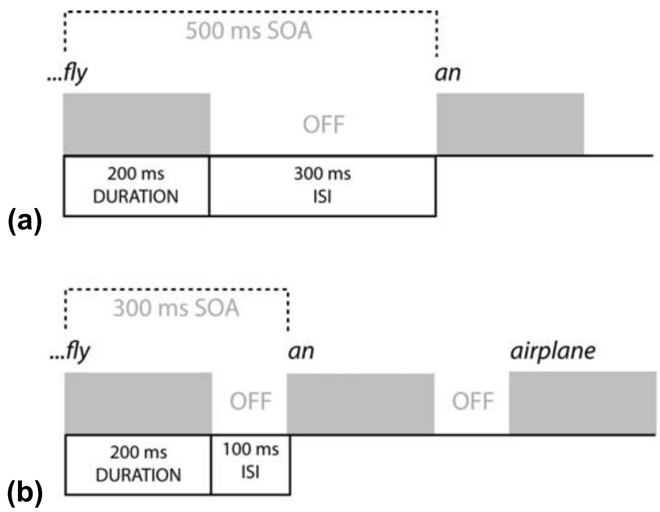


Figure 3.1. Stimulus onset asynchrony (SOA) differences between (a) Experiment 1 and (b) the current study.

3.3.2. Participants

Thirty-two volunteers (24 women) participated in the ERP experiment for course credit or for cash. All were right-handed, native English speakers with normal or corrected-to-normal vision, between 18-35 years (mean, 21.7 years). Thirteen participants reported a left-handed parent or sibling.

3.3.3. Offline measures

Twenty-two of the original 32 ERP participants were administered additional offline tests in a separate testing session. Because this testing was conducted at a later date, 10 of the original ERP participants were no longer available to complete this session. Tests were administered by a single individual in a private testing room free of distractions. Testing included an author recognition task (ART), which entailed identification of real authors' names from a list including equal numbers of foils. (The ART questionnaire is included in **Appendix B**.) Our version of this questionnaire was modeled, in part, after one proposed by Stanovich & West (1989). The ART data was collected as a rough index of amount of reading experience and exposure to print media. Of the 22 participants, ART scores ranged from 18-73% correctly identified authors, with mean author identification of 28%.

Participants also performed a reading comprehension task which included two essays with multiple choice questions. Number of correct answers was recorded as well as overall essay reading times. This test was conducted with the idea that potential variability in our ERP results might in part be attributable to differences between participants' natural reading rates. Of the 22 participants, average reading times (over both essays) ranged from 2.54-5.66 words per second, with a mean reading time of 3.9 words/second.

In the following Results section, any analyses reported based upon ART or reading comprehension scores included only those 22 participants who were administered these tests offline.

3.3.4. Electroencephalographic recording parameters

These were the same as for Experiment 1 of this thesis.

3.3.5. Data analysis

Trials contaminated by eye movements, excessive muscle activity, or amplifier blocking were rejected offline before averaging – on average, 7% of articles and 10% of nouns. Data with excessive blinks were corrected using a spatial filter algorithm. A digital band-pass filter set from 0.2 to 15 Hz was used on all data to reduce high frequency noise.

Data were re-referenced offline to the algebraic sum of left and right mastoids and averaged for each experimental condition, time-locked to the target article and noun onsets.

Where appropriate, Huynh-Feldt (HF) epsilon correction was performed, and reported throughout are the original degrees of freedom and the HF-corrected *p*-values.

3.4. Behavioral Results

Comprehension accuracy was calculated for the yes/no probe questions following the RSVP presented sentence stimuli. Participants correctly answered an average of 95.3% (range: 88% to 100%) of the questions, indicating that they were attending to and comprehending the experimental sentences during the recording session.

3.5. ERP results

For both the articles and nouns, we performed discrete (ANOVA) as well as continuous (correlation) analyses.

3.5.1. Articles

For the articles we analyzed ERP mean amplitude in the N400 time window relative to article cloze values and relative to the cloze of the upcoming target nouns.

3.5.1.1. Discrete analyses

Conducting analyses similar to those performed for Experiment 1 of this thesis, we began by examining ERP effects at the articles. An omnibus ANOVA with 2 levels of article cloze (<50%[LO], ≥50%[HI]) X 26 electrodes revealed that between 200-500 ms, there was no main effect of article cloze in the article N400 time window, [$F(1,31) = .99$, $p = .33$, n.s.], nor was there a significant interaction of cloze by electrode site (**Figure 3.2**). Similarly, when the same analysis was conducted except sorting the articles based on low (<50%) and high (≥50%) *upcoming noun* cloze (**Figure 3.3**), there was also not a main effect of noun cloze

[$F(1,31) = .43$, $p = .52$, n.s.], nor an interaction with electrode site. We further explored possible cloze effects at the article by contrasting only the extreme ends of article cloze ranges (0-9.9% versus 90-100%), which revealed no main effects of article cloze [$F(1,31) = 1.29$, $p = .26$, n.s.] nor interaction with electrode. However, using extreme ends of the upcoming noun cloze range (**Figure 3.4**) revealed a significant effect of noun cloze [$F(1,31) = 4.94$, $p = .03$] and an interaction of cloze with electrode site [$F(25,775) = 2.73$, $pHF = .03$], which indicated that ERPs to the articles preceding the least expected nouns were more negative than those preceding the most expected nouns, particularly at medial scalp locations (**Figure 3.5**).

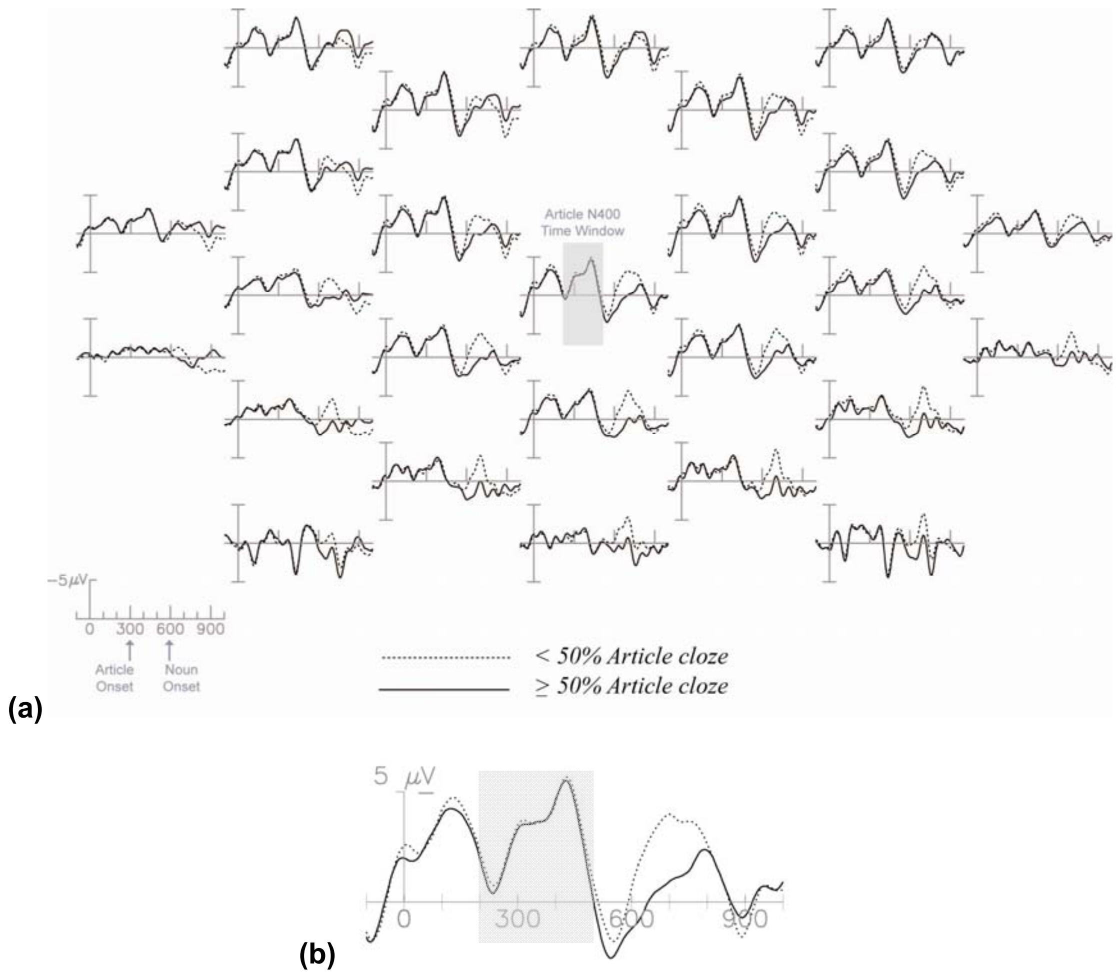
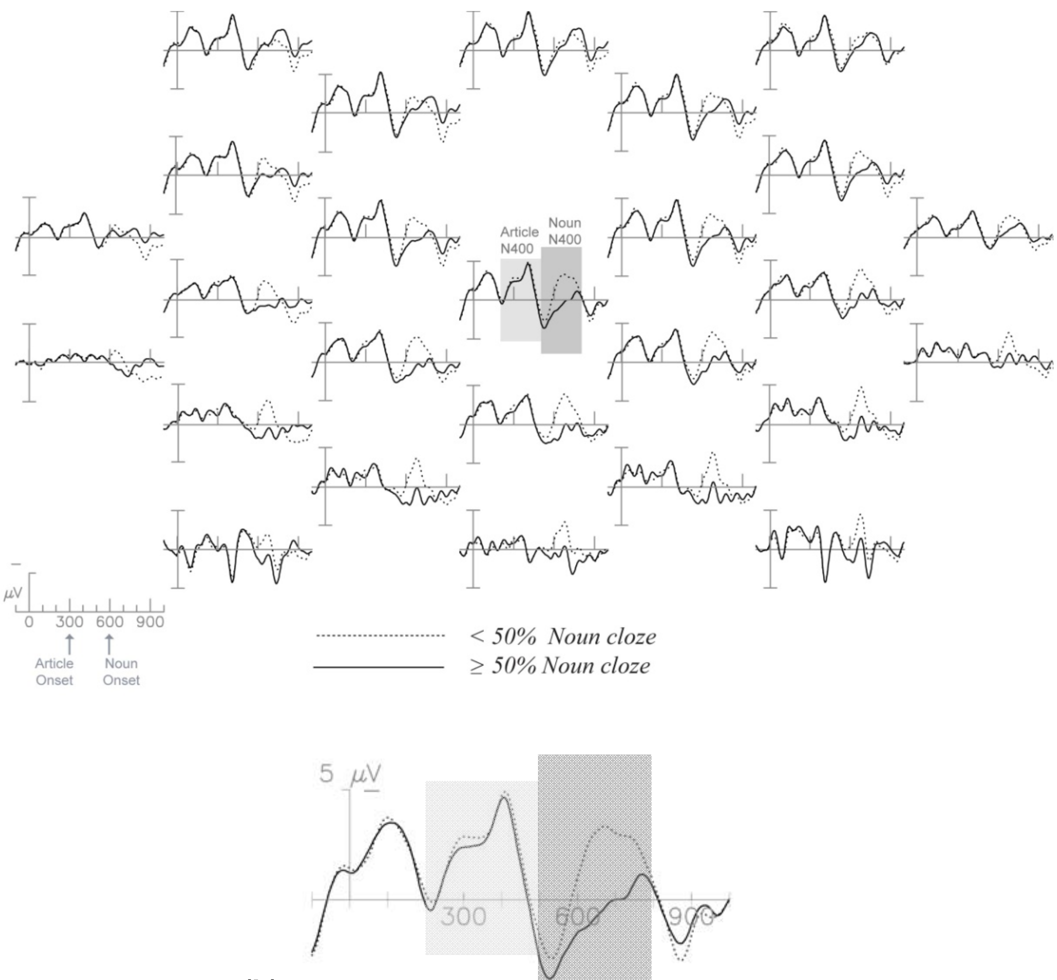


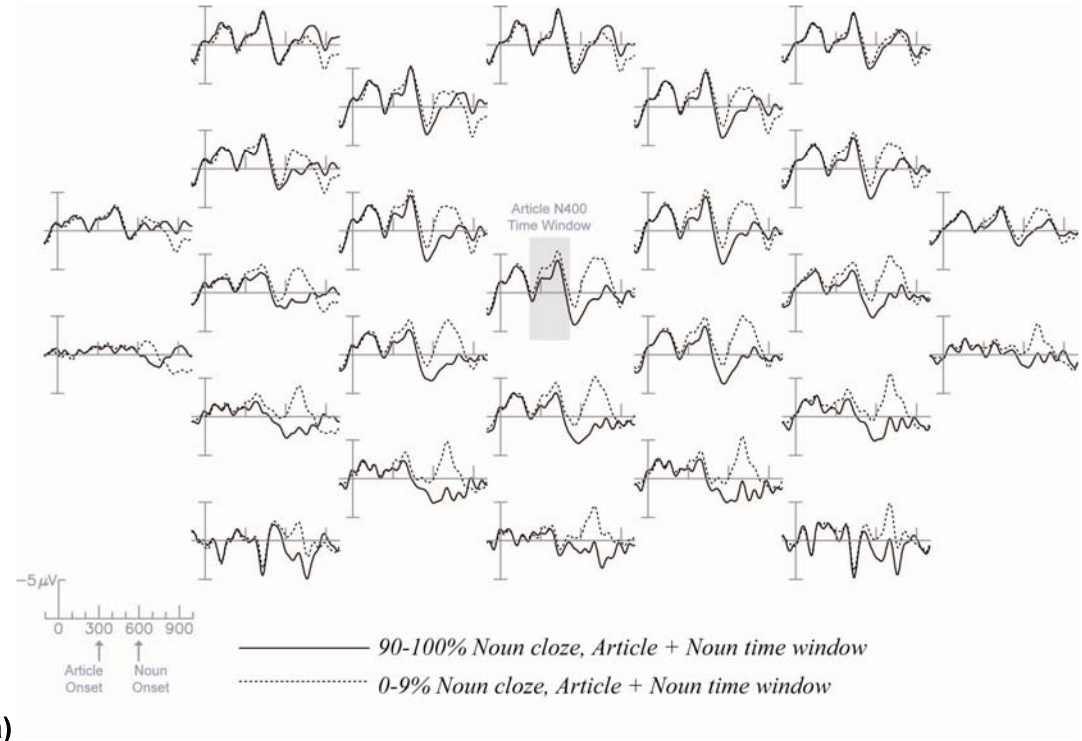
Figure 3.2. Article + noun time window sorted by high (≥50%) and low (<50%) article cloze, with article N400 window highlighted, (a) over all 26 electrodes and (b) at the vertex channel.



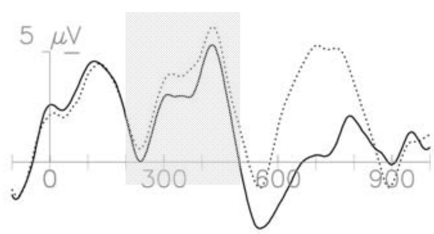
(a)

(b)

Figure 3.3. Article + noun time window sorted by high (≥50%) and low (<50%) noun cloze, with article and noun N400 windows highlighted, (a) over all 26 electrodes and (b) at the vertex channel.



(a)



(b)

Figure 3.4. Article + noun time window sorted by extreme ends of noun cloze range, very high (90-100% noun cloze) versus very low (<10% noun cloze), (a) over all 26 electrodes and (b) at the vertex channel, with article N400 time window highlighted.

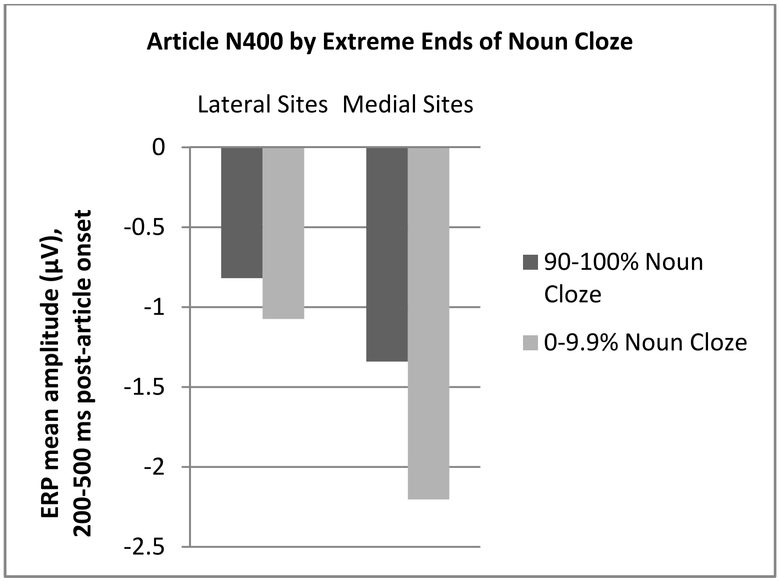


Figure 3.5. Article N400s sorted by highest (90-100%) and lowest (0-9.9%) noun cloze, cloze interacting with laterality.

3.5.1.2. Article Correlations

3.5.1.2.1. Article mean amplitude with Article cloze

Correlations of cloze probability with article ERP mean amplitude (200-500 ms) were also conducted. Possible results for the correlational analysis in the N400 time window are outlined in **Figure 3.6**. When article cloze probability was used (**Figure 3.7a**), we did not observe any significant pattern of increasing N400 amplitude (negativity) with decreasing cloze (as would be indicated by positive *r*-values). Instead, and only at left, frontal scalp sites was there a pattern of increased positivity with decreasing cloze (reflected in the negative *r*-values), though even these maximal correlations did not reach significance. These results contrast sharply to the N400-like correlations observed for the same analysis when words were presented with a 500 ms SOA, as in our original study, Experiment 1 (**Figure 3.7b**).

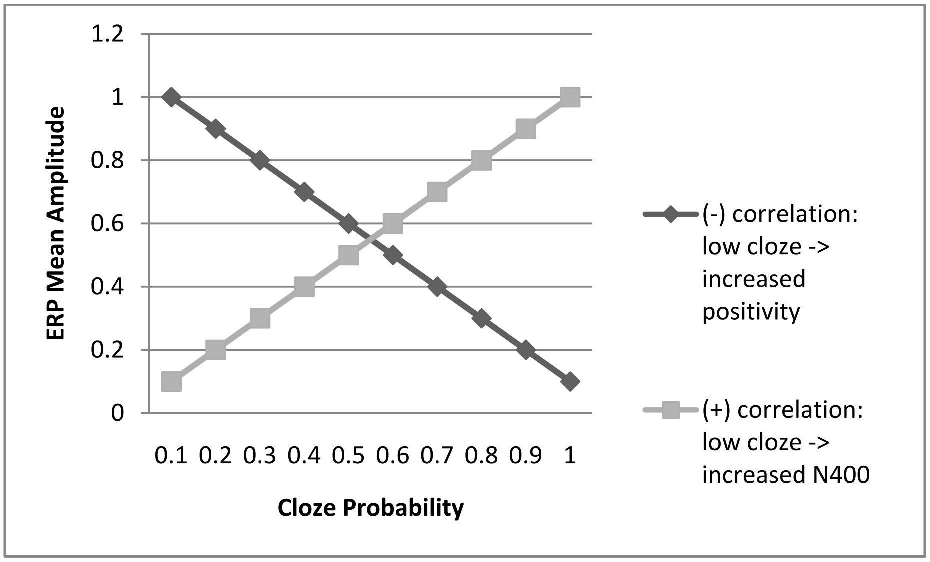


Figure 3.6. Possible correlation mappings of cloze probability with mean amplitude in the N400 time window.

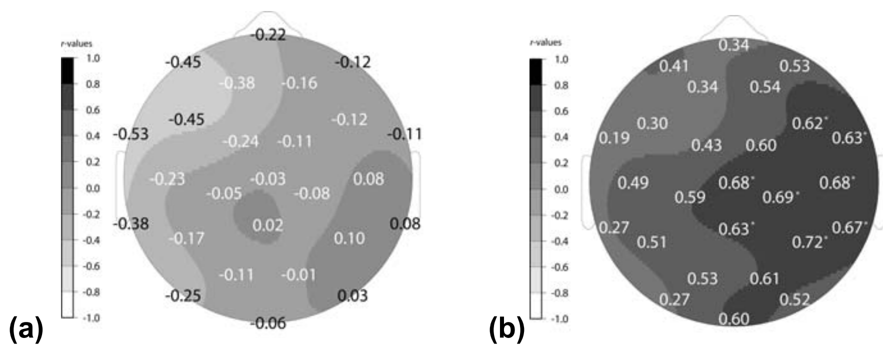
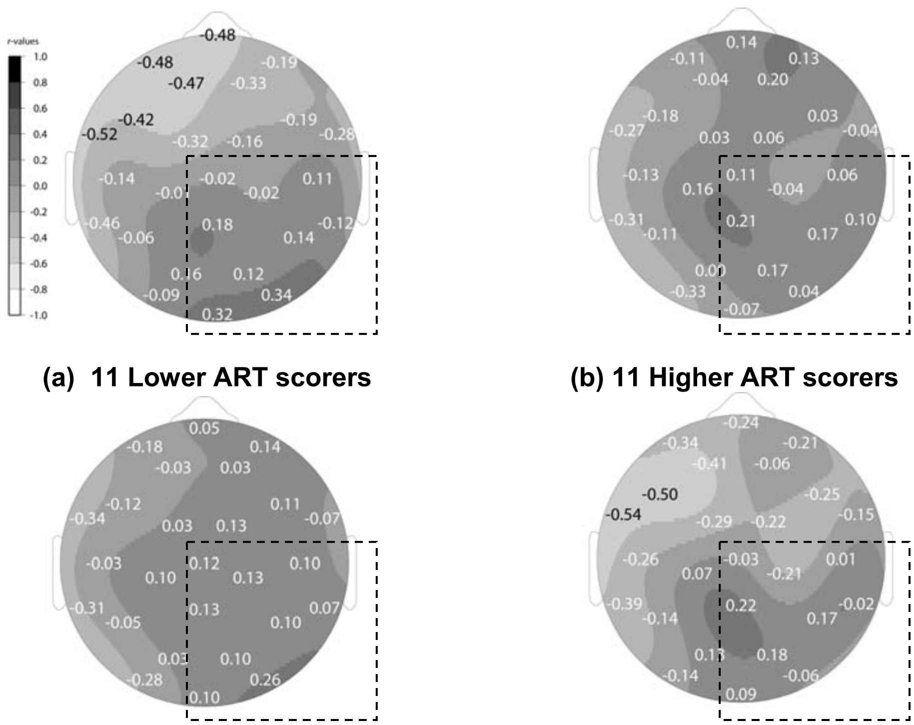


Figure 3.7. 200-500 ms post article onset. Article mean amplitude correlated with article cloze from (a) current study using 300 ms SOA, and (b) original study (Experiment 1) using 500 ms SOA. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

We also explored possible group differences in effects of cloze probability correlated with article mean amplitude (**Figure 3.8**). In our correlations of article cloze with article mean amplitude (200-500 ms), neither comparisons of high versus low ART scorers nor comparisons of faster versus slower readers revealed any substantive correlation differences (though not tested statistically) over the scalp areas where N400 effects are typically maximal (i.e., over the right posterior scalp quadrant). In particular, across participant groups there was a similar lack of an article N400/article cloze effect (an N400-consistent correlation relationship would be indicated by positive *r*-values).



(a) 11 Lower ART scorers

(b) 11 Higher ART scorers

(c) 11 Slower readers

(d) 11 Faster readers

Figure 3.8. 200-500 ms post article onset. Article mean amplitude correlated with article cloze for different participant groups. Positive *r*-values indicate increasing ERP negativity with decreasing cloze (a more N400-like pattern). Negative *r*-values indicate increasing ERP positivity with decreasing cloze. Scalp areas highlighted by boxes indicate primary regions for comparison, where N400 effects are typically maximal, though no statistically significant correlations ($p \leq .05$) were present in these analyses.

3.5.1.2.2. Article mean amplitude with Noun cloze

We also analyzed the article N400 time window by correlating article mean amplitude with the cloze probability of the upcoming nouns. This analysis (**Figure 3.9**) revealed consistently low *r*-values (non-significant) across the scalp, though reflecting a more N400-like pattern of increased negativity with decreasing cloze, compared to the correlations of article mean amplitude with article cloze (**Figure 3.7a**).

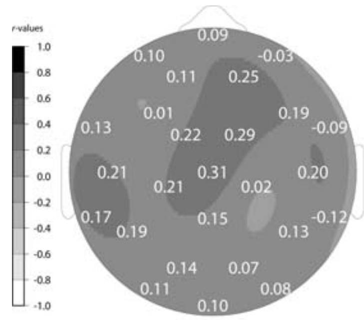
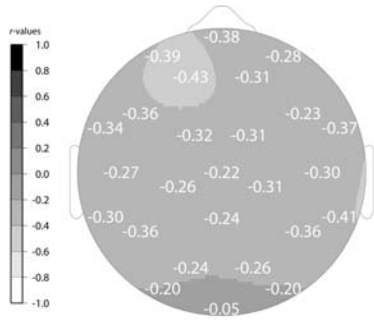
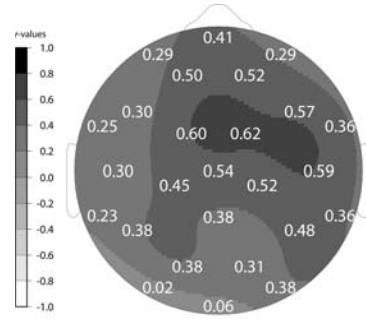


Figure 3.9. 200-500 ms post Article onset. Article mean amplitude correlated with noun cloze. No statistically significant (p ≤ .05) correlations were present in this analysis.

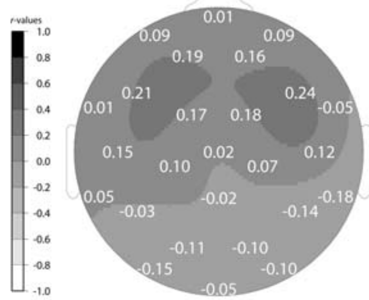
Again, we examined possible group differences in patterns of correlations of article mean amplitude, this time with the cloze of the upcoming nouns (Figure 3.10). The general correlation patterns for both slower and faster readers (panels 3.10c and 3.10d) were similar: low (non-significant) *r*-values across the scalp. However, when participants were classified by their ART scores, the high ART scorers (panel 3.10b) exhibited more N400-like correlations than participants with low ART scores (panel 3.10a), with maximal correlation values over midline and right central-anterior electrodes. These correlation strengths and the distribution pattern are similar to those observed for correlations of article mean amplitude with article cloze in Experiment 1, Figure 3.7b.



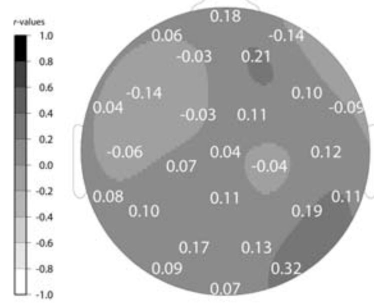
(a) 11 Lower ART scores



(b) 11 Higher ART scores



(c) 11 Slower readers



(d) 11 Faster readers

Figure 3.10. 200-500 ms post article onset, Article mean amplitude correlated with upcoming noun cloze for different participant groups. Positive *r*-values indicate increasing ERP negativity with decreasing cloze (a more N400-like pattern). Negative *r*-values indicate increasing ERP positivity with decreasing cloze. Only the Higher ART scorers showed correlation values that approached, though did not reach, statistical significance (p ≤ .05).

3.5.2. Nouns

Visual inspection of the noun waveforms suggested that in addition to an effect of cloze within the noun 200-500 ms time window, this effect reversed over an extended time window following the N400. This pattern of results suggested an LP effect similar to that observed in Experiment 1, thus motivating analyses in both the N400 and the LP time windows.

3.5.2.1. N400 time window (200-500 ms)

3.5.2.1.3. Discrete analyses

Mean amplitude ERPs were also calculated for the nouns. See **Figure 3.3**. Again, contrasting 2 levels of noun cloze (<50%[LO], ≥50%[HI]) X 26 electrodes, an omnibus ANOVA revealed a significant N400 effect between 200-500 post noun onset [$F(1,31) = 34.87$, $p < .0001$], with less expected nouns exhibiting a mean amplitude of 0.18 μV relative to that for more expected nouns (1.38 μV). There was also a significant interaction of noun cloze X electrode [$F(25,775) = 18.47$, $p < .0001$] which revealed the largest N400 effects to be at posterior, medial scalp sites. A pattern of increased negativity to less expected relative to expected nouns was present over all but the left, lateral, frontal sites, where this pattern reversed to one of increased positivity to less expected nouns.

3.5.2.1.4. Correlations

Correlations of noun cloze probability and noun mean amplitude indicated a pattern similar to that for the same analysis in Experiment 1: widely-distributed high N400/cloze correlation values maximal at right, posterior scalp locations (**Figure 3.11**). However, over left, anterior scalp areas a reversal in the direction of the correlation (increasing positivity with decreasing cloze) hinted that the frontal positivity (LP) observed in Experiment 1 may also be present with speeded presentation.

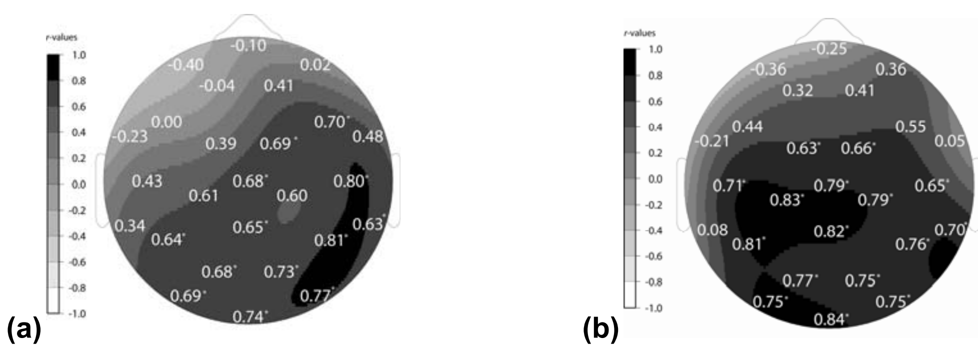


Figure 3.11. 200-500 ms post noun onset, N400 time window. Noun mean amplitude correlated with noun cloze from (a) current study using 300 ms SOA, and (b) original study (Experiment 1) using 500 ms SOA. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

3.5.2.2. Late positivity time window (500-900 ms)

To investigate a possible late frontal positivity effect, we performed both discrete and correlational analyses of noun cloze and noun mean amplitude in a later time window (500-900 ms). An ANOVA similar to the one performed for the N400 time window revealed that in the LP time window, ERPs to less expected nouns were more positive than those to more expected nouns [$F(1,31) = 5.00$, $p = .03$]. See **Figure 3.12**. This effect interacted with electrode site [$F(25,775) = 8.29$, $pHF < .0001$], revealing effect sizes that gradually increased from posterior to anterior sites, which were more left lateralized than right, and which were largest at medial sites.

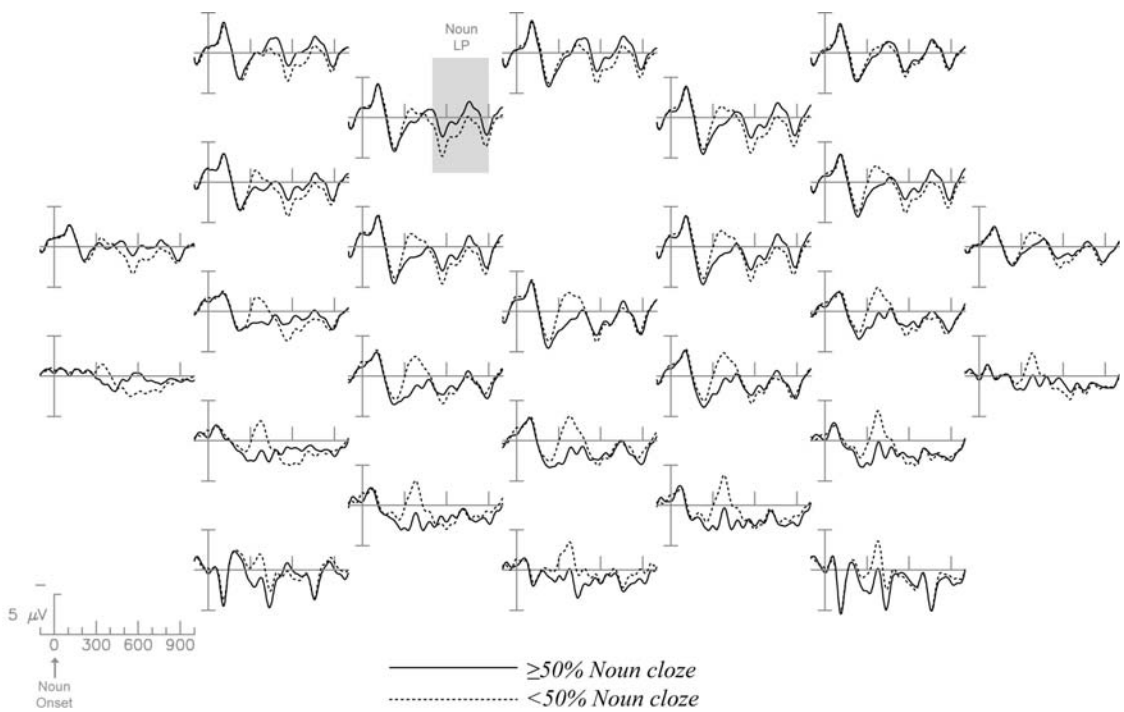


Figure 3.12. Nouns sorted by high (≥50%) and low (<50%) noun cloze. The LP time window (500-900 ms) is highlighted.

Correlation analyses also indicated that similar to – but to an even greater extent than – the original study, there was an increase in LP amplitude with decreasing noun cloze, with maximal LP correlations present at anterior electrode channels (**Figure 3.13**).

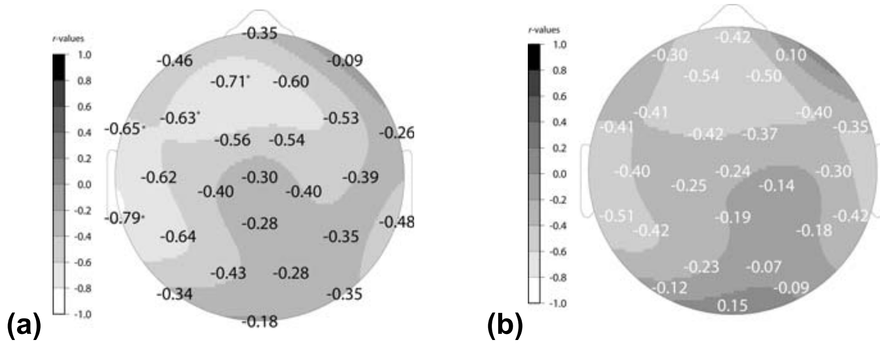


Figure 3.13. 500-900 ms post noun onset, LP time window. Noun mean amplitude correlated with noun cloze from (a) current study using 300 ms SOA, and (b) original study (Experiment 1) using 500 ms SOA. Statistically significant correlations (p ≤ .05) are indicated with asterisks (*).

3.6. Discussion

The current study was designed to test whether the ERP effects evidencing linguistic prediction observed in Experiment 1 continued to be present when the rate of input was increased from 2 words per second to 3.3 words per second – conditions more consistent with the upper range of average reading rates for college students (Lewandowski et al., 2003). A potential argument against the Experiment 1 findings is that an RSVP rate of 2 words per second may have allowed readers “spare time” to predict and that our finding of a cloze graded N400 at prenominal articles as evidence of prediction was only obtainable under non-naturalistic conditions. Prediction studies by Wicha et al. (2003a,b; 2004) and van Berkum and colleagues (van Berkum et al., 2005; Otten et al., 2007) argue against this idea in that they observed pre-noun target prediction effects (though binary, not graded) resulting from auditorily presented natural speech stimuli. Also contradicting this proposal are results from ERP studies which show that varying RSVP rates (within limits) does not seem to implicate alternate processing mechanisms (Kutas, 1987; Gunter, Jackson, & Mulder, 1992). Nonetheless, the potential concern about the limits of graded, probabilistic prediction is best addressed empirically.

At the slower presentation rate, we observed correlations of pre-nominal article mean amplitude with article cloze probability in the article N400 time window which were similar to – though not as high as – the correlations of noun N400 mean amplitude and noun cloze in that study (as well as nouns in other studies, e.g., Kutas & Hillyard, 1984). At the faster RSVP rate, however, we did not observe the same graded effect of article N400 with article cloze. Instead, there was an N400-like main effect of expectancy between 200-500 ms across all 32 participants when the articles were sorted according to the cloze probability of their respective following nouns. Articles preceding nouns with the lowest noun cloze probabilities (low cloze [0-10%, mean cloze 1%] showed greater ERP negativity than those preceding nouns with the highest cloze [90-100%, mean cloze 97%]). In addition, though we did not observe a graded effect of article N400 amplitude with upcoming noun cloze across all 32 subjects, for a subset of participants – those exhibiting high ART scores – there was a strong correlation pattern of article N400 amplitude with noun cloze probability, with correlation strengths and a topographical distribution similar to those observed for the article N400/article cloze correlations in Experiment 1.

The binary nature of the noun-cloze based ERP effect at the article in the current study is similar to ERP prediction effects observed in previous studies of prediction by van Berkum et al. (2005) and Wicha et al (2004), in which ERP amplitude associated with prenominal gender-marked words varied according to whether the following target nouns were either highly expected or unexpected (>75% cloze probability for expected nouns, mean cloze 86%, and 2% mean cloze for unexpected nouns in van Berkum et al., 2005; mean expected noun cloze probability ranging from 65-100%, 80% mean cloze for Wicha et al.) In both of these experiments, the cloze probability utilized for categorizing the prenominal words as expected or unexpected was that of the target nouns (not that of the prenominal words themselves). These conditions are similar to the ones under which prediction effects were observed in the current experiment.

On the other hand, the graded article N400 prediction effect correlated with the cloze of the following noun observed within a subset of participants in our study (the 11 showing the highest ART scores), is a novel finding. And though it differs from the prediction effect observed at the article when a slower presentation rate was used (hinging now on the cloze of the upcoming noun, rather than the article itself), it nonetheless offers some intriguing evidence for probabilistic pre-activation (at least in some readers), and indicates that comprehenders are not simply waiting until linguistic input has been received in order to begin integrating it with preceding context – even when the individual words of a sentence must be processed relatively quickly.

A comparison of ERP effects at the target nouns across slower and faster presentation rates, however, revealed highly similar patterns in both the N400 (200-500 ms) and LP (500-900 ms) time windows. Similar to the 500 ms SOA, for the 300 ms SOA nouns N400 amplitude decreased as a function of increasing cloze probability in the typical manner as measured through both discrete (ANOVA) and continuous (correlational) analyses, with widespread but maximal posterior N400 patterns across the scalp. Both discrete and correlational effects in the noun LP time window were also similar across both SOAs: increased ERP positivity as a function of decreasing noun cloze, with maximal effects at left anterior electrode sites. In fact, the strength of the LP/noun cloze correlations was even greater at the faster presentation rate (maximum r = -.79) compared to the slower SOA (maximum r = -.54).

To summarize our ERP findings, at the nouns there was no substantial difference in cloze effects as a function of rate of presentation. The ERP effects at the articles, however, differed between studies, with article N400 amplitude varying as a function of article cloze probability at the slower SOA, but as a function of the upcoming noun's cloze probability at the faster SOA. At the faster rate, these article N400 effects were graded only over a subset of participants with high ART scores, though there was a main effect at the extreme ends of noun cloze probability over all 32 participants.

3.6.1. Article ERP prediction effects

3.6.1.1. Comparison of two SOAs

The results of the current study suggest that while the prenominal prediction effects at the article for the faster presentation rate are non-identical to those observed at the slower SOA, they still offer evidence for anticipatory language processing. We will provide some possible explanations for what we consider to be the three major comparison points between the article prediction effects of the two studies.

The first difference relates to the article prediction effect being associated with article cloze at a slower presentation rate, but with noun cloze at the faster SOA. Results at the slower SOA had suggested that across highly and less constraining contexts, the word form for the most likely upcoming noun had already been pre-activated by the time the article was presented, and via this noun pre-activation, the preceding article was projected and evaluated as being more or less easy to integrate. We suggest that at the faster presentation rate, the reduced processing time may limit the degree to which the articles themselves are anticipated, since determining the prediction-consistent article requires mapping from the anticipated nouns. The fast SOA article N400/noun cloze effect may thus reflect the parser simply "checking" whether the article is consistent with expected noun, rather the evaluating the probability of the article itself. Our lack of an article N400/*article* cloze effect is also consistent with the idea that a specific indefinite article is likely pre-activated at a point temporally closer to the noun it precedes. For instance, in the sentence '*The day was breezy so the boy went outside to fly…*', preactivation of *kite* might strengthen throughout the sentence, though it seems less likely that *kite's* preceding determiner (*a*) would experience similar incremental pre-activation. More likely, '*a*' gets triggered closer to the syntactic point at which the pre-activated noun is about to occur (i.e., when a noun phrase has been signaled with the presentation of the verb *fly*). So while the slower rate may have afforded enough time from the onset of *fly* to activate/select the lexical form of *kite* and also enough time for the prenominal determiner to have been pre-activated, the faster rate may have only allowed enough time for preactivation of the noun itself.

A second difference between different SOAs is that at the faster presentation rate, over all 32 participants, the prediction effect at the article was binary (at the extreme ends of noun cloze) rather than graded. It may be that in highly constraining contexts, targets may be preactivated well before they are encountered in the input, with a number of varied and obvious constraint sources (e.g., semantic features, lexical association, event knowledge, frequency, syntactic order, etc.) pointing toward a single, highly probable continuation. For these sentences, a "checking" strategy is sufficient to reveal an ERP effect at the prenominal article because the prediction has been well formed. On the other hand, in a less constraining context the parser might rely more heavily on even the most subtle sources of constraint (e.g., phonological form of prenominal indefinite articles), with pre-activation levels getting an additional "boost" closer to the point the item will most likely be encountered. In these cases, an article ERP prediction effect may not be observable because rather than simply (dis)confirming a prediction for an upcoming noun, the article itself is playing a larger role in shaping that prediction at the "last minute". For instance, for a less constraining sentence such as '*Sue had wanted to go to Tim's birthday party but she was still waiting for…'*, where norming revealed that individuals supplied a variety of continuations (e.g., *an invitation, a ride, her date*, etc.), the parser may rely more on the article information (*a* or *an*) to help narrow the possibilities for what comes next, rather than being used to check for consistency with a highly expected noun.

The third notable comparison between the studies using faster and slower presentation rates is that a graded prediction effect observed over all participants at a slower presentation rate was observed over only a subset of participants at the faster presentation rate. At the faster rate, only those participants exhibiting higher ART scores showed a clear correlation pattern of increased article N400 amplitude with decreasing noun cloze. These results suggest that at least some comprehenders – the more facile or experienced readers – are able to make better use of even limited constraint to preactivate upcoming information well before the input is received.

3.6.1.2. Implication of article results for prediction

We suggest, then, that the results from the current study are consistent with the information value of the prenominal determiners varying depending on contextual constraint. In highly constraining sentence contexts, when upcoming nouns might be preactivated well before they are encountered, it is possible that the article is processed more at the level of "pattern checking". If all cues (e.g., event knowledge, frequency, semantic memory, etc.) point to a continuation being highly likely, then the article itself would not be projected to have as much information value. In these sentences, then, it may be that the unfolding context provides enough constraint even prior to the presentation of the article such that prediction of the lexical form of the upcoming noun has reached a maximum. The noun lexeme information is thus available to evaluate the fit of the preceding article.

However, for less constraining contexts, where a number of nouns may be preactivated to varying degrees, perhaps the article itself carries much more weight in shaping predictions; in other words, instead of simply "monitoring" the article for consistency with the expected noun, the article itself has a higher informational value. Consequently, noun preactivation levels may be modulated in closer temporal proximity to receipt of the target in the less constraining sentence contexts, which may explain why there is less of a graded cloze effect for the faster SOA.

In sum, it seems likely that the article may play slightly different roles in highly and less constraining contexts. For high constraint sentences the article may serve more as confirmatory information, as verification that the pre-activated noun is indeed about to appear. In the mid/lower constraint sentences the article may be used more to constrain the choices of the upcoming noun if no one "competitor" dominates up until that point, with the article playing a larger part in strengthening or weakening preactivation levels. With a slower input rate the parser exhibits a sensitivity to both of these potential roles of the article, while with limited processing time at the fast SOA, only the parsers' sensitivity to more well-formed predictions may have been detectable.

3.6.1.3. Potential methodological explanation of article results

Although we have offered a potential explanation for why graded expectancy effects were not observed across all brainwave participants at the faster presentation rate, we believe there may be another, less theory-driven explanation for this pattern of results. This alternative proposal focuses on the issue of overlapping ERP components and the possibility that this factor may have contributed to obscuring a prediction-related ERP effect at the article. Indeed, though the noun N400 and LP main effect sizes from the original study (Experiment 1) were quite large (on the order of 1-2 μV), the article N400 main effect from this study was already quite small: in fact, a cloze-based article ERP difference was revealed only in the correlations, as the binary effect was not significant when contrasting articles with cloze probabilities above and below 50%. At the faster rate, with an SOA of 300 ms, the peak of the N400 to the prenominal articles directly overlaps with the frontal N1 (50-150 ms post onset) to the following target noun. The frontal N1 component has been linked to the allocation of visuospatial attention and has been proposed to reflect the output of a capacity-limited attentional system (Clark & Hillyard, 1996; Mangun, Hillyard & Luck, 1993). And in language processing, Federmeier & Kutas (1999) observed reduced frontal N1's to expected targets in highly constraining contexts. In relation to our own data, we cannot rule out the possibility that participants may have unconsciously picked up on the cue value of pre-target indefinite articles as a signal to the target nouns (which although sometimes highly expected, at other times were less expected, thus possibly contributing to the articles being more salient), even though the pre-target articles were not the only ones in the sentence stimuli. If the prenominal articles somehow served to signal the brain that it should prepare the attentional system to perhaps devote more resources for what was about to come – regardless of the cloze of the article or noun – then this may have led to an increase in noun frontal N1 amplitude that was concurrent with the N400 to the preceding article. In effect, this may have dampened any (relatively subtle) article prediction effects. Though this overlap of components could have been an issue in the current experiment due to the SOA timing of the stimuli, it would not have occurred at the slower presentation rate. This overlap would also have been avoided by presenting the sentences as natural speech stimuli in the auditory modality (similar to van Berkum, 2005 and Wicha et al., 2003b), where spoken word lengths would have varied to a greater extent. Van Berkum et al. also circumvented the overlapping component problem by always including one or more intervening words between their prenominal gender-marked adjectives and the target nouns (e.g., *a big*$_{gender\text{-}marked}$ *but rather unobtrusive painting*$_{gender}$).

3.6.2. Noun ERP effects

3.6.2.1. Noun N400

Consistent with results from our original study, at the shortened SOA, noun cloze probability modulated noun N400 mean amplitude in the typical way. As noun cloze increased, noun N400 amplitude decreased, with the largest effects over central posterior scalp electrodes. Also consistent with our earlier findings, at the fast SOA the posterior N400 pattern gave way to an increased positivity to less expected nouns at left frontal sites, hinting that – like in the original study – the LP effect may begin as early as the N400 time window. Like previous studies that have manipulated visual presentation rate of sentence stimuli, then, the noun N400 effect did not seem to be affected by the (within normal limits) increase in presentation rate.

3.6.2.2. Noun late positivity (LP)

Our findings related to the LP held across presentation rate. In addition to frontal positivities of this sort having been described infrequently in the language ERP literature, to our knowledge this is the first study to show that such an effect persists at an increased presentation rate. In fact, the noun cloze/LP mean amplitude correlations in the present study appeared to be even stronger than those elicited at the slower presentation rate of Experiment 1.

Though speculative, we proposed in the Discussion section of Experiment 1 that the LP may be related to some form of prediction violation, with low cloze nouns eliciting more positive ERPs than high cloze nouns in the time window following the N400. In contrast to our N400 prediction effects at the prenominal article, which were not so robust as to be unaffected by rate of presentation, a strong LP effect consistent over presentation rate offers some preliminary support for the idea that additional cognitive processing is required when predictions are not substantiated by physical input. Additional testing of this ERP effect is undoubtedly needed to shed light on its functional nature, in particular to explore whether this postlexical effect is indeed a consequence of prediction, and if so what type of prediction (e.g., syntactic or more general anticipatory processing). We will explore these questions further in Chapter 4 of this thesis.

3.7. Conclusions

Although the current evidence for linguistic prediction at a rate of 3.3 words per second is non-identical to that for the materials presented at a rate of 2 words per second, our results continue to support a model of predictive language processing. In light of the differences between Experiments 1 and 2, it seems clear that an auditorily-presented, natural speech version of the current experiment would offer the most ecologically valid way to test for graded prediction effects at increased input rates, while avoiding some of the methodological pitfalls of (very) rapid SVP. This study is certainly in the works. Yet even under the current design, the parser indicated a sensitivity to the consistency of pre-target words with an anticipated upcoming input stream, particularly in cases where there was a single, well-fitting continuation to the sentence. And although not all participants showed a graded, probabilistic prediction response, a subset did; namely, those who might be considered to be more experienced readers. Finally, our finding of a late positive ERP effect that increases in a graded manner with decreasing cloze probability – an effect similarly observed at the slower SOA – may turn out to be linked to the consequences of mispredicting. In combination with the pre-target predictive ERP findings, such evidence would go a long way toward fleshing out a model of anticipatory language processing, and ultimately toward establishing the benefits and costs of linguistic preactivation.

3.8. Appendix B: Offline Author Recognition Test (ART)

CODE:__________ DATE:__________

Author Recognition Questionnaire

Below you will see a list of 80 names. Some of the people in the list are popular writers (of books, magazine articles, and/or newspaper columns) and some are not. You are to read the names and put a check mark next to the names of those individuals who you know to be writers. Do not guess, but only check those who you know to be writers. Remember, some of the names are people who are not popular writers, so guessing can easily be detected.

1. Marilyn Jager Adams ___
2. Richard Allington ___
3. Donna Alvermann ___
4. Maya Angelou ___
5. Isaac Asimov ___
6. Kathryn Au ___
7. Rebecca Barr ___
8. Isabel Beck ___
9. Judy Blume ___
10. Erma Bombeck ___
11. Hilda Borko ___
12. Bertram Bruce ___
13. P. E. Bryant ___
14. Robert Calfee ___
15. Barbara Cartland ___
16. Carlos Castaneda ___
17. Jeanne Chall ___
18. Tom Clancy ___
19. Arthur C. Clarke ___
20. James Clavell ___
21. Theodore Clymer ___
22. Max Coltheart ___
23. Stephen Coonts ___
24. Priscilla Drum ___
25. Gerald Duffy ___
26. Dolores Durkin ___
27. Robert Dykstra ___
28. John Elkins ___
29. Roger Farr ___
30. Ian Fleming ___
31. James Flood ___
32. Dick Francis ___
33. Ruth Garner ___
34. Jack Goody ___
35. Stephen J. Gould ___
36. Michael Graves ___
37. Andrew Greeley ___
38. John Guthrie ___
39. David Halberstam ___
40. Alex Haley ___
41. Jane Hansen ___
42. Shirley Brice Heath ___
43. Frank Herbert ___
44. S. E. Hinton ___
45. John Jakes ___
46. Beau Fly Jones ___
47. Erica Jong ___
48. Michael Kamil ___
49. Stephen King ___
50. Dean Koontz ___
51. Judith Krantz ___
52. Louis L'Amour ___
53. Isabelle Liberman ___
54. Robert Ludlum ___
55. George McConkie ___
56. James Michener ___
57. P. David Pearson ___
58. Susanna W. Pflaum ___
59. Sylvia Porter ___
60. Keith Rayner ___
61. Nelson Rodriguez-Trujillo ___
62. Nancy Roser ___
63. S. Jay Samuels ___
64. Sidney Sheldon ___
65. Danielle Steel ___
66. Barbara M. Taylor ___
67. Paul Theroux ___
68. Robert Tierney ___
69. Alvin Toffler ___
70. J. R. R. Tolkien ___
71. Barbara Tuchman ___
72. John Updike ___
73. Leon Uris ___
74. Richard Venezky ___
75. Irving Wallace ___
76. Alice Walker ___
77. Joseph Wambaugh ___
78. Samuel Weintraub ___
79. Tom Wolfe ___
80. Bob Woodward ___

C	
M	
F	
%	

3.9. References

Clark, V.P. & Hillyard, S.A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, 8(5), 387-402.

Federmeier, K.D. & Kutas, M. (1999). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Gunter, T.C., Jackson, J.L. & Mulder, G. (1992). An electrophysiological study of semantic processing in young and middle-aged academics. *Psychophysiology* 29(1), 38-54.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161-163.

Kutas, M. (1987). Event-related brain potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. In *Current Trends in Event-Related Brain Potential Research, EEG* Suppl. 40 (eds. Johnson, R., Rohrbaugh, J.W. & Parasuraman, R.) 406-411 (Elsevier, Amsterdam).

Lewandowski, L.J., Codding, R.S., Kleinmann, A.E. & Tucker, K.L. (2003). Assessment of reading rate in postsecondary students. *Journal of Psychoeducational Assessment*, 21(2), 136-144.

Mangun, G.R., Hillyard, S.A. & Luck, S.J. (1993). Electrocortical substrates of visual selective attention. In D. E. Meyer, & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*. (219-243). Cambridge, MA, US: The MIT Press.

Otten, M., Nieuwland, M. & Van Berkum, J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, 8(1), 89.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.

Reynolds, M.E. & Givens, J. (2001). Presentation rate in comprehension of natural and synthesized speech. *Perceptual and Motor Skills*, 92(3, Pt 2), 958-968.

Stanovich, K.E. & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24(4), 402-433.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

Wicha, N.Y., Moreno, E.M. & Kutas, M. (2003a). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483-508.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003b). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

CHAPTER 4.
EXPERIMENTS 3A,B:
EVIDENCE OF A PROCESSING COST FOR PREDICTION IN LANGUAGE COMPREHENSION

4.1. Abstract

In Experiment 1 of this thesis – a sentence processing study that tested for lexical prediction by examining ERPs to English indefinite articles (*a*, *an*) preceding more and less expected noun continuations – in addition to an anticipated cloze-related noun N400 effect, our analyses revealed a prolonged, late frontal positivity to unexpected nouns (e.g., *check* in *Dale was very sorry and knew he owed Mary a check...*, where *an apology* is the expected continuation). In line with theories proposing that a somewhat similar late ERP component, the P600, reflects aspects of syntactic processing, we hypothesized that our late positivity (LP) might reflect the violation of a contextual expectancy for a part of speech: namely, an adjective (e.g., *sincere*), which would be pre-activated upon the receipt of an unexpected article (e.g., *a*). Key to the development of this hypothesis were the results from an off-line sentence norming study (Experiment 3A of this chapter) in which sentence contexts were normed with both types of indefinite articles. These findings confirmed that offline, different contexts elicited adjective+expected noun responses to varying degrees. In the face of an unexpected indefinite article, interjecting an adjective allows the parser to salvage the most contextually expected noun (e.g., *apology*), especially when it is difficult to activate an alternative continuation quickly. We propose that when the parser adopts this strategy during online comprehension, additional processing may then be required when it encounters an unexpected noun, for integration and resolution of the "syntactic surprise" at not receiving an adjective.

To test this hypothesis, in ERP Experiment 3B participants read sentence contexts like the one above, with four possible continuations: (1) expected article/expected noun (*an apology*), (2) unexpected article/unexpected noun (*a check*), (3) expected article/unexpected noun (*an answer*), and a filler condition (4) unexpected article/adjective/expected noun (*a sincere apology*). If the LP previously observed to *check* is due to anticipating adjectives instead of nouns after unexpected articles are presented, then *answer* (Condition 3) should elicit only an N400, with no LP, because it was preceded by the prediction-consistent article, and therefore there would be no need to propose an intermediate adjective; conversely, *check* (Condition 2) should elicit both an N400 and an LP.

Contrary to our predictions, along with increased posterior N400 amplitude (correlated with cloze probability) we observed prolonged LPs to *both* unexpected noun conditions (*check* and *answer*), indicating that the effect was not isolated to the condition associated with "syntactic surprise". Further analyses of our ERP results revealed that whether unexpected nouns were preceded by expected or by unexpected articles, the LP turned out to be highly correlated with the expected (but not presented) noun's cloze probability – a measure equivalent to sentential constraint. We argue that these results provide further support for an anticipatory model of language processing, by crucially identifying a "cost" (and a graded one, at that) to mispredicting. Additionally, our results indicate a double dissociation between ERP indices of cloze probability (N400) and contextual constraint violation (LP). Finally, we offer a possible explanation for the LP observed in our study which is compatible with other recent findings of P600s/LPs to semantic manipulations, though further work will be needed to determine the precise functional correlate of the component.

4.2. Introduction

4.2.1. Background

At various levels of cognition, it has been argued that top-down, associative processing may be the brain's "default" mode of operation (e.g., Bar, 2007; Engel, Fries, & Singer, 2001). With biological continuity as a guide, it is not unreasonable to consider prediction in the linguistic domain. Indeed, one can imagine that anticipatory processing could confer some advantages, for instance in ease of processing, potentially freeing up resources, allowing for response preparation in conversational situations or allowing extra time for contemplation when input rates are out of a comprehender's control. These benefits of prediction could be part and parcel of how the language comprehension system operates, coming "for free" as a consequence of how the brain organizes and processes language information. The interesting questions relating to prediction, then, might best be considered in terms of what an anticipatory processor might reveal about how the brain utilizes contextual elements to construct meaning on the fly.

Although anticipatory language comprehension constitutes a plausible parsing strategy, for decades the dominant view has been that comprehension proceeds primarily in a bottom-up fashion (e.g., Forster, 1989; Marslen-Wilson, 1989; Norris, 1994), and that prediction – with its potential for error – is an unlikely and inefficient strategy. This has been a contentious issue in the sentence processing literature, with debates centering on whether comprehenders use context to activate in advance upcoming items or features of items, or whether the effects of context come into play only after input has been received and identified.

In recent years, though, event related brain potential and eye-tracking researchers have found ample evidence for anticipatory processing occurring at various linguistic levels (e.g., syntactic, semantic, phonological, and lexical) during language comprehension (e.g., Altmann & Kamide,1999; Kamide, Altmann & Haywood, 2003; Kamide, Scheepers, Altmann, & Crocker, 2002; Federmeier & Kutas, 1999; Wicha, Moreno & Kutas, 2003a, 2004; Wicha, Bates, Moreno & Kutas, 2003b; DeLong, Urbach & Kutas, 2005; Federmeier, 2007; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Koornneef & van Berkum, 2006; Otten, Nieuwland & Van Berkum, 2007). These on-line methodologies allow predictive effects to be observed both preceding and following more and less anticipated word targets. So as evidence continues to mount for predictive language comprehension, a parallel question is whether there are consequences to receiving input that has *not* been pre-activated to some degree and which would presumably alter the direction of the mental sentence representation constructed to that point.

This brings up what might be one of the barriers to broader acceptance of predictive views; that is, if the parser functions predictively, then there should be some reflection of a cost for un- or less expected continuations when the parser gets it wrong. With early naming time studies finding no evidence for such costs (e.g., Stanovich & West, 1981 and 1983), predictive processing models did not find a wide following. After all, even the most efficient parser will not always anticipate correctly, perhaps leading to some reduction of available cognitive resources by way of backtracking, reanalysis, or revision when sentences continue in some unexpected way.

The idea of "cost" in sentence processing is nothing new, but at least until very recently this idea has mostly been discussed with respect to syntactic processing. Within language ERP research, the P600 has been considered by many (e.g., Hagoort, Brown & Groothusen, 1993; Friederici, 1995; Münte, Matzke, and Johannes, 1997; Osterhout, Holcomb, & Swinney, 1994) to be a component reflecting reanalysis of structural relations: the more difficult or ambiguous the syntactic parse, the larger the P600. Though some might be inclined to similarly view the N400 as reflecting the cost of semantic integration, this would not be accurate. The N400 component is thought to index the degree of semantic fit of an item within a particular context; for instance, every word in a sentence elicits an N400, but based on particular factors – a major one being the degree to which surrounding context has facilitated an item – some items experience a *reduction* in amplitude, reflecting the degree to which an item's integration has been eased, rather than an increase in N400 amplitude reflecting a semantic processing cost. In fact, the N400 has been shown to be insensitive to the degree of constraint violation, with low cloze continuations to both highly and not-at-all constraining contexts exhibiting similar amplitude N400s (Kutas & Hillyard, 1984). So with regard to semantic processing, an ERP finding reflecting a consequence to mispredicting has remained elusive.

In contrast, terms often associated with the P600 – such as syntactic "reanalysis", "revision" or "repair" – by definition entail the parser already having committed to a particular analysis of the sentence structure, for instance to the assignment of thematic roles, and then being forced by the input to re-evaluate the previous parse (e.g., *The broker persuaded to sell the stock was sent to jail*, from Osterhout & Holcomb, 1992). However it is rare, if at all, that these "re-" processes are framed in terms of a prediction gone wrong; rather, a majority of psycholinguists would suggest that processing costs reflected by P600s indicate an incompatibility of the input with the incremental parse of the preceding material up until the point of the ambiguity/ungrammaticality. So although P600 effects in response to syntactically ambiguous, complex or ungrammatical continuations are not *incompatible* with a predictive strategy – i.e., the P600 indexing a violation of (more or less syntactic) prediction – the effect has generally not been framed as such in the literature. Though there have been a few recent studies reporting effects of syntactic prediction (aka top-down parsing strategies) – e.g., Staub & Clifton (2006); Altmann & Kamide (1999); Kamide, Altmann, & Haywood (2003) – the eye-tracking methodology used for these studies cannot address the potential association of the P600 as an electrophysiological signature of a syntactic prediction violation. Nor do the scopes of these studies speak to the even larger question of whether syntactic prediction and its potential violations are just isolated instances of a more general predictive strategy.

For over a decade and a half now, many language ERP-ers have drawn a clear distinction between the cognitive processes indexed by the N400 and P600 ERP components: the N400 reflects semantic processing, while the P600 indexes syntactic processing. In particular, the specificity of the P600 as being a "syntactic" component is consistent with a more modular, domain-specific model of how the brain's language parser interprets linguistic input. While to date there has been little evidence disputing the N400 as being associated exclusively with semantic processing (but see Deacon, Dynowska, Ritter & Grose-Fifer, 2004, for a different view), theories about the P600 have been less consistent, with some researchers contending that that processes indexed by the P600 are not syntax specific (e.g., Coulson, King & Kutas, 1998). Though the P600 does indeed exhibit an amplitude and latency sensitive to syntactic violations, ambiguities, and complexities, claiming that the P600 is functionally correlated exclusively with syntactic processing is likely an overgeneralization. Nearly since its discovery in the early 1990s (Osterhout & Holcomb, 1992), it has been debated whether the P600 is a specific index of syntactic processing, and more recently (within the past five years or so), even stronger challenges to this view have emerged. In ERP sentence and discourse comprehension research, a number of research groups have converged on findings of LPs in situations where N400s were projected, to so-called "semantic anomalies" (see **Table 4.1**). Though in the strictest sense many of the eliciting sentence stimuli are considered grammatical, it is unclear whether the brain's parser would interpret them this way, or whether they might be processed more like grammatical errors. For instance, in some of the experimental manipulations a word order or morpheme change would make the sentence stimuli semantically-thematically repairable: e.g., '*The javelin has/was by the athletes thrown*', (translated from the original Dutch), or Hoeks, Stowe & Doedens, 2004; '*The hearty meal was devouring/devoured*', Kim & Osterhout, 2005. In many cases these LPs occurred instead of expected N400s; in others, the LPs occurred in conjunction with N400s. With theories of what these LPs index ranging from ideas about semantic attraction overriding syntax (Kim & Osterhout, 2005), to the monitoring of conflicts of plausibility-based heuristics and syntactic analyses (Kolk, Chwilla, van Herten & Oor, 2003; van Herten, Chwilla & Kolk, 2006; van Herten, Kolk & Chwilla, 2005), to thematic role assignment costs (Hoeks et al., 2004), to extended combinatory analysis of parallel thematic and syntactic streams (Kuperberg, 2007), to animacy violations (Nieuwland & van Berkum, 2005), to reanalysis of thematic structure (Bornkessel, Schlesewsky & Friederici, 2002 and 2003), a clear picture is still evolving.

Table 4.1. Findings of P600s (LPs) to "semantic" violations (independent of or in addition to N400s).

Researchers/ Study	Experimental manipulation	Example stimuli (Underlined words represent locale of ERP effect, bold indicates the eliciting condition)	ERP results	LP Latency	LP Scalp Distribution	Task	Sentence Target Position	Language
Geyer, Holcomb, Kuperberg & Perlmutter (2006)	Task effect: Plausibility judgment task vs. passive reading	*Tyler cancelled the subscription because she was not entirely satisfied.* ***Tyler cancelled the tongue…***	N400 for both conditions, P600 for plausibility judgment only	Estimated visually to be from 500/600-1100 ms	Estimated visually to be more posterior	Comprehension question vs. plausibility judgment	Sentence medial, clause final	English
Hoeks, Stowe & Doedens (2004)	Lexico-semantic fit/message level constraint	*The javelin was by the athletes thrown.* ***The javelin has the athletes thrown.***	No N400, P600	600-1100 ms	centro-parietal, slightly more leftish	Plausibility judgment	Sentence final	Dutch
Hoeks, Stowe & Doedens (2004)	Semantic verb-argument violation with no verb-argument association	*The javelin was by the athletes thrown.* ***The javelin has the athletes summarized.***	N400 + P600					
Kim & Osterhout, (2005)	Semantic verb-argument violations	*The hearty meal was devoured…* ***The hearty meal was devouring…***	No N400, P600	600-900 ms	posterior	Acceptability judgment	Sentence medial	English.
Kolk, Chwilla, van Herten & Oor (2003)	Semantic reversals	*The mice that from the cat fled…* ***The cat that from the mice fled …***	No N400, P600	650-850 ms measured for significance, though effects extend to 1000 ms	centro-parietal	Acceptability judgment in one version/Comp questions only in another	Sentence medial	Dutch
Kolk, Chwilla, van Herten & Oor (2003)	Task effect: Plausibility vs. reading for comprehension	*The trees that in the park stood* ***The trees that park played***	Plausibility judgment = P600, Passive reading = N400					

Table 4.1 continued

Kuperberg, Caplan, Sitnikova, Eddy, Holcomb (2006)	Thematically-semantically attracted but not repairable/reversible	*Every morning at breakfast the boys would eat…* ***To make good documentaries cameras must interview…***	No N400, P600	500-900 ms	widespread but maximal at posterior	Sense/acceptability judgment	Sentence medial	English
Kuperberg, Kreher, Sitnikova, Caplan & Holcomb (2007)	Semantic verb-argument violation with no verb-argument association	*Every morning at breakfast the boys would eat…* ***Every morning at breakfast the eggs would plant…***	No N400, P600	550-850 ms	widespread but maximal at left posterior	Sense/acceptability judgment	Sentence medial	English
Kuperberg, Sitnikova, Caplan, Holcomb (2003)	Semantic-thematic relationship violation (animacy)	*Every morning at breakfast the boys would eat…* *Every morning at breakfast the boys would plant…* ***Every morning at breakfast the eggs would eat…***	No N400, P600	500-800 ms	centro-parietal, more marked at posterior than at anterior	Sense/acceptability judgment	Sentence medial	English
Nieuwland & Van Berkum (2005)	Animacy violation within semantically congruent discourse context	Following a semantically congruent discourse context: *…the woman told the tourist…* ***…the woman told the suitcase…***	No N400, P600	emerging around 500 ms, significant from 700-1300 ms	larger at posterior	Read for comprehension (no additional task)	Discourse medial	Dutch
Osterhout, Holcomb & Swinney (1994)	Sucategorization bias to disambiguating words	*The doctor believed the patient was lying.* ***The doctor charged the patient was lying.***	No N400, P600	500-800 ms	widespread, larger at posterior	Acceptability judgment	Sentence medial	English
Sitnikova, Kuperberg, Holcomb (2003)	Contextually appropriate and inappropriate objects in video film clips of common activities	Silent movie clip of man getting ready to shave, reaches for a razor and begins shaving. **Silent movie clip of man getting ready to shave, reaches for a rolling pin and begins "shaving".**	N400 + P600	600-900 ms	posterior	Sense judgment (is it a scenario from everyday life?)	NA	NA
Van de Meerendonk, Kolk, Vissers & Chwilla (in press)	Plausible vs. highly implausible vs. mildly implausible sentence continuations	*The eye consisting of among other things a pupil, iris and retina…* *The eye consisting of among other things a pupil, iris and eyebrow…* ***The eye consisting of among other things a pupil, iris and sticker…***	N400 + P600	500-800 ms	central-posterior	Sentence recognition task between blocks	Sentence medial	Dutch
Van Herten, Chwilla & Kolk (2006)	Verb-argument semantic association	*John saw that the elephants pushed over the trees.* *John saw that the elephants pruned the trees.* ***John saw that the elephants caressed the trees.***	*pruned* = P600, *caressed* = N400	600-800 ms	largest over central and posterior sites	Comprehension questions	Sentence medial	Dutch
Vissers, Kolk, van de Meerendonk & Chwilla (2008)	Violations of sentence truth-values in picture-sentence matching task	Following either a compatible or **incompatible** picture presentation: *The triangle stands below the square.*	P600, Early negative effect	500-700 ms	posterior	Mismatch detection	NA	Dutch

Table 4.1 continued

Moreno, Federmeier & Kutas (2002)	Spanish /English bilinguals. Low cloze words completing English sentence fragments and idioms, and to unexpected switches into Spanish.	Straight: *The driver of the speeding car was given a* ***citation/multa*** *(ticket).* Idiomatic: *The truth hit me like a ton of* ***stones /ladrillos*** *(bricks).*	N400 to lexical switches, LAN and large posterior positivity (450–850 m) for code switches, and late frontal positivity for both	650-850 ms	anterior	Experimental final idiom completion task	Sentence final	English /Spanish
Coulson & Van Petten (2007)	Low cloze literal sentence endings (compared to high cloze literal endings)	*The doctor told him his headaches were due to* ***hypochondria****.* (Lateralized presentation)	N400 + LP (only with RVF presentation)	600-900 ms	anterior	Report lateralized word and T/F comp questions	Sentence final	English
Coulson & Wu (2005)	Probes following joke (related) or straight (unrelated) sentence continuations	*CRAZY* following *Everyone had so much fun jumping into the swimming pool, we decided to put in a little water/platform* (Lateralized presentation)	N400 + LP (larger for RVF than LVF presentation)	700-900 ms	anterior	Relatedness judgment to probe	Sentence final	English
Swick, Kutas & Knight (1998)	Incongruent vs. congruent sentence endings	unavailable	N400 + LP	600-900 ms	unavailable	--	--	--
DeLong (Experiment 1, doctoral dissertation)	Expected vs. unexpected sentence endings	*The day was breezy so the boy went outside to fly a kite/****airplane…***	N400 + LP	500-1200 ms	anterior, left bias	Comprehension questions	Sentence medial	English
Federmeier, Wlotko, de Ochoa & Kutas (2007)	Expected vs. unexpected endings in strong vs. weak constraint sentences	*The children went outside to play. The children went outside to look. Joy was too frightened to move.* ***Joy was too frightened to look.***	N400 + LP	500-900 ms	anterior	Experiment final recognition test	Sentence final	English

4.2.2. The current studies

This issue of LPs in response to semantic manipulations was highlighted by a surprising ERP result from Experiment 1. In this study, the main finding was one of probabilistic pre-activation of specific word forms in sentence contexts as evidenced by a graded N400-like negativity to pre-nominal indefinite articles (*a*, *an*) of varying cloze probability. For the contextually unexpected nouns in that study (e.g., *check* in *Dale was sorry for what he said to Mary. He knew he owed her a check*..., where *an apology* is expected), in addition to a posterior N400 inversely correlated with cloze probability, an enhanced LP to low cloze relative to high cloze nouns was observed between 500-1200 ms, largest at frontal scalp sites (see **Figure 4.1**).

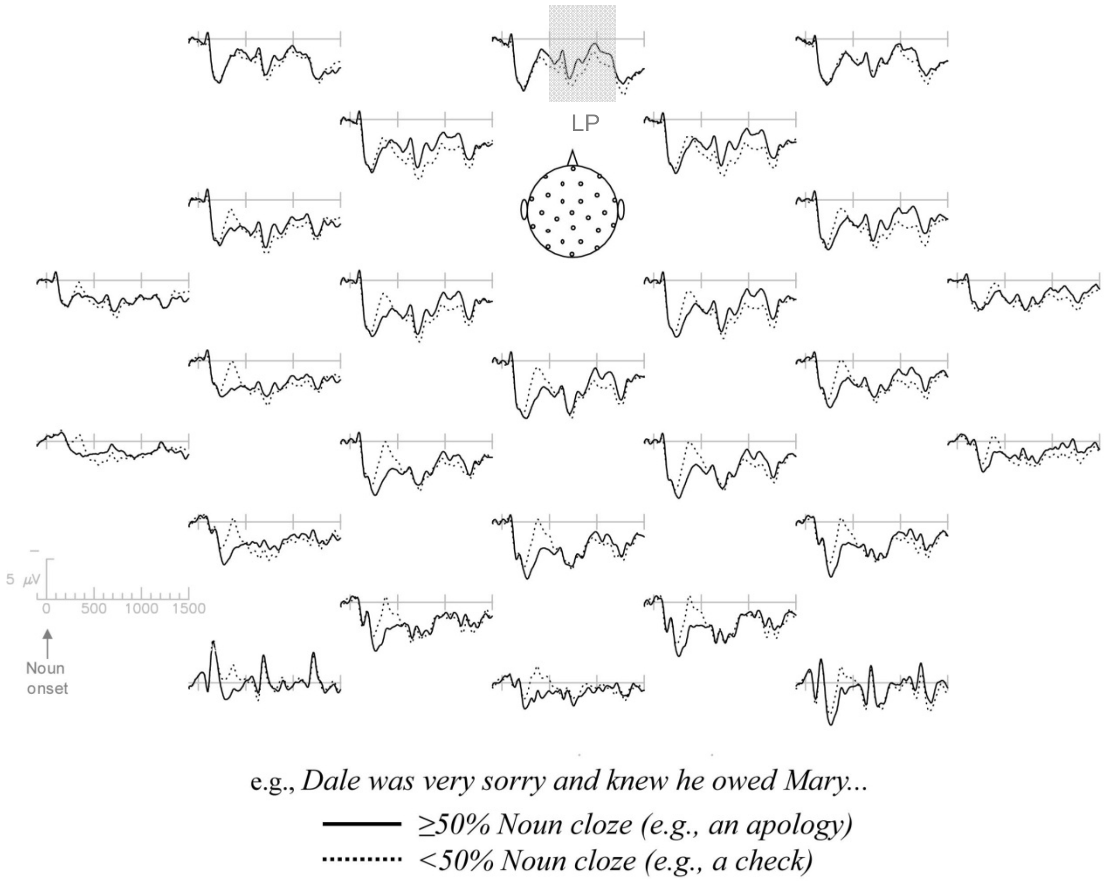


Figure 4.1. ERP plot of contextually more versus less expected nouns, over all 26 scalp electrodes, from Experiment 1 of this thesis. LP time window is highlighted.

In addition to traditional P600 findings, this reminded us of some of the recent findings of "semantic" P600s outlined in **Table 4.1**. Our manipulation, though, was semantic in the strictest sense. Our unexpected finding of the post-N400 positivity prompted us to examine our stimuli more closely, and hypothesize possible explanations for this effect. The first possibility that occurred to us, based on the long history of LPs being associated with syntactic manipulations, was that there was something about our stimuli that was leading individuals to alter their syntactic expectations. For instance, in the preceding stimulus example, *apology* is the highest cloze probability noun prior to receiving the unexpected indefinite article *a*; however, the unexpected *a* forces the parser to pursue an alternative continuation. We considered that the positivity might be elicited by the violation of a contextual expectancy for a particular part of speech, namely an adjective (for example, *sincere* or *heartfelt*), which the parser activates once the unexpected indefinite article *a* appears. This strategy would allow the parser to salvage the most contextually activated noun (*apology*), presumably because it is difficult to readily activate an alternative continuation. However, once the low cloze noun (*check*) appears, additional processing would then be required not only to integrate the unexpected noun into the sentence structure, but also to resolve the syntactic surprise of not receiving an adjective.

To follow up on this proposal, we first wanted to determine whether comprehenders were indeed salvaging their noun expectations by means of intermediate adjectives offline. To test this, we began by performing an in-depth analysis of the cloze probability norming results for the stimuli used in Experiment 1. These findings are reported as Experiment 3A of the current chapter. In particular, we examined results of the stimulus norming for the truncated sentence contexts that included the unexpected indefinite articles. Across these contexts, we were curious whether participants continued the sentences, to varying degrees, with consonant or vowel-consistent adjectives followed by the original highly expected noun. If this were the case, then these results would indicate that adjective insertion at least held off-line, thus elevating the possibility that the LP might be associated with such a strategy during on-line processing.

In the present ERP study (Experiment 3B), then, we explore the hypothesis that the LP following the N400 to low cloze probability (unexpected) nouns reflects "syntactic surprise" at receiving a contextually unexpected part of speech. To test this idea, we used sentence stimuli with indefinite article+noun continuations ranging in cloze probability. The sentences also modulated the contextual expectancy for adjectives by way of the indefinite articles. In this way, we could determine whether there were ERP effects that reflect violation of syntactic constraint when adjectives are expected but not presented. In particular, we utilized the two experimental conditions from Experiment 1: sentence stimuli that continued with either a more expected (high cloze indefinite article + high cloze noun) or a less expected (low cloze indefinite article + low cloze noun) target pair (e.g., *Dale was very sorry and knew he owed her an apology/a check for what he had done*.) Additionally, we included a third experimental condition, comprised of a high cloze indefinite article+low cloze noun (e.g., *an answer*), which was designed to test whether the noun frontal positivity observed to *check* in our original study was due to expecting an adjective (e.g. *sincere*, *heartfelt*) instead of a noun after the low cloze probability article (*a*) is presented. The idea behind contrasting the *check* and *answer* conditions was that if the subject is forced by the low cloze article (*a*) to revise their original expectation (*apology*), then we would expect to observe an N400+LP to *check* relative to *apology*, but for *answer* we would expect to observe only an N400. This is because for *answer* individuals would have received the expected article and would never have had a reason to switch their expectation from noun to adjective in the first place. Thus, if the syntactic expectancy hypothesis is supported, then we would expect to observe large N400 amplitudes to both unexpected noun continuation types (*check* and *answer*), but an LP only for the unexpected article condition (*check*), an ERP pattern idealized in **Figure 4.2**. Finally, we included a fourth condition which was comprised of the following: the unexpected article + a semantically congruous adjective + the expected noun (e.g., *a sincere apology*). Our reasoning for doing this was to ensure that there were indeed instances where adjectives were actually presented, as well as to supplement the proportion of more contextually expected noun continuations. Regarding analysis of this "adjective" condition, we make no experimental predictions due to the differences in word class and features between this condition and the other noun conditions.

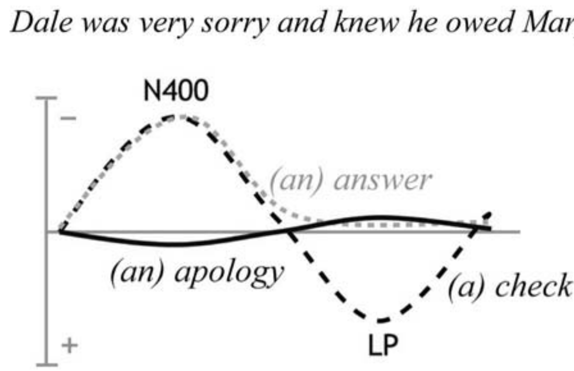


Figure 4.2. Hypothetical data patterns under our syntactic expectancy hypothesis.

If our syntactic expectancy hypothesis is upheld, and an LP is noted only to unexpected nouns like *check* where a prediction-inconsistent article forces a (syntactic) modification of expectancies, then this would be compatible with the idea that comprehenders were forming expectations for a particular word class, and that their expectations were violated when they did not receive that word class. This syntactic expectancy account would also be in line with long-held views of the P600 as being a component associated with the increased processing costs at points of syntactic ambiguity or unexpectedness at a structural level. If, however, LPs are observed to *both* unexpected noun conditions, then a syntactic explanation of the ERP pattern does not fit, as the prediction compatibility/incompatibility of preceding article would not seem to be contributing to the effect. Finding no difference between the two unexpected noun conditions would be more compatible with the LP being associated with a more general cost to receiving an unexpected item when the context constrains for something else.

In addition to examining the potential LP effects, we also expect to replicate both the target indefinite article and noun N400 findings from Experiment 1 – that is, the ERP negativity between 200-500 ms that increases as a function of the items' cloze probabilities. We also plan to examine earlier components of the ERP; namely, the frontal N1 (50-150 ms) and P2 (150-250 ms) components. Though such early ERPs are typically associated with more perceptual analysis, these components have on occasion been shown to be sensitive to semantic manipulations. The frontal N1 has been associated with the allocation of visuospatial attention (e.g., Clark & Hillyard, 1996; Mangun, Hillyard, & Luck, 1993), with enhanced N1s being elicited in response to target stimuli in attended (as opposed to unattended) locations. Federmeier & Kutas (2001) observed reduced frontal N1s to expected picture endings in highly constraining versus less constraining sentence contexts, arguing that these context effects may play a role in reducing attentional load. The P2 is a frontally distributed positive-going ERP peaking around 200 ms, and is part of the normal electrophysiological response to visual stimuli. Amplitude modulations of the P2 have been linked to the detection and analysis of visual features in paradigms such as selective attention tasks (Hillyard & Münte, 1984; Luck & Hillyard, 1994). In visual hemifield studies, where frontal P2 effects to language stimuli seem to be commonly observed, Federmeier & Kutas (2001) have noted enhanced P2s to expected relative to unexpected pictorial sentence endings for RVF presentation, which they interpret as top-down LH sentence processing allowing for enhanced visual feature extraction from expected targets. And Federmeier, Mai & Kutas (2005) argue that frontal P2s to highly but not weakly constrained word targets in sentence stimuli for RVF presentation are compatible with prediction effects, where the LH is better able to make use of context and "prepare" for upcoming input. Additionally, Wlotko & Federmeier (2007) showed that for RVF presentation, the frontal P2 was sensitive to constraint, regardless of the target word's cloze probability. Though the frontal P2 findings mentioned here were all for lateralized presentation, these results hint that an investigation of this component for our centrally presented data may also be warranted, given our somewhat similar experimental manipulations and goals.

EXPERIMENT 3A: A NORMING STUDY

As a first step toward investigating our proposal that "syntactic surprise" may contribute to the LP effect described in Experiment 1 of this thesis, we wanted to more systematically examine the kind of responses comprehenders supplied offline in the cloze probability norming task of that study. In particular, we were interested in finding out how participants continued sentence contexts in which the unexpected indefinite articles were supplied. Please refer to the Methods section of Experiment 1 for details of the materials and cloze probability norming procedure.

For the stimulus norming of items used in the Experiment 1, cloze probabilities for the target articles and nouns were first determined by norming the sentence contexts without any indefinite article supplied, e.g., '*The bakery did not accept credit cards so Peter had to write…*'. The results for the norming without articles indicated that there were wide ranges of cloze probabilities for both the articles and the nouns (**Figures 4.3a,b**). However, given the potential for indefinite articles to alter the degree of preactivation of a particular item within a sentence context, it was informative to test how individuals completed the sentences off-line when the articles were provided. Each sentence context was thus normed with both article types (*a* and *an*), albeit by different groups of participants. In most cases (as determined by norming the contexts without the indefinite articles supplied) one article type was more expected, i.e., had a higher cloze probability, than the other type. There were 60 volunteers total for the norming with articles, with each context being completed by 30 participants. Each participant saw only one version of each sentence context.

(a)

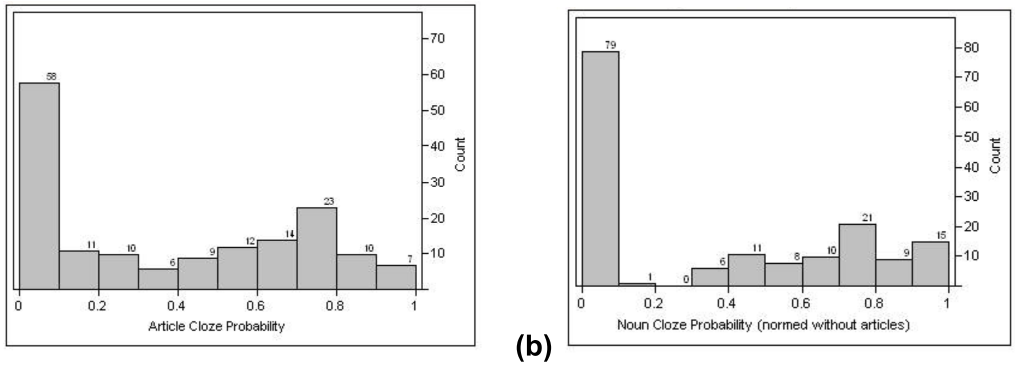


Figure 4.3. Cloze probability distributions of (a) articles and (b) nouns normed without articles in Experiment 1.

4.3. Norming with expected articles

For norming with expected articles, participants supplied what they considered to be the best completion for sentences such as '*The bakery did not accept credit cards so Peter had to write a…*' (expected noun is *check*). The goal to norming with the expected articles was to determine if the indefinite articles altered sentence constraint and changed the noun cloze probabilities in any significant way. The results from the original norming without articles had indicated that on average, mean expected noun cloze probability was 0.70. When normed with the expected articles, noun cloze probability increased by an average of 14% to 0.84. We concluded from these results that the indefinite articles generally increased sentence constraint and raised the noun cloze probability due to the narrowing of possible continuations. For instance, including the indefinite articles eliminates the possibility for encountering mass nouns such as *water* or *furniture* and it also restricts the parts of speech that may potentially follow (only a noun, adjective or adverb are possible). In addition, this finding suggests that when individuals encounter the target indefinite article there might be some degree of local (i.e., last minute) strengthening of noun expectations.

4.4. Norming with unexpected articles

Along with examining how the expected articles altered sentence constraint and noun cloze probabilities, we also wanted to determine what happens off-line when the article information disconfirms a noun prediction. To test this, the same truncated sentence contexts were normed except this time with the unexpected instead of expected articles supplied. The goal here was to examine how individuals "recovered" from receiving unexpected articles. For instance, when presented with the following sentence, '*Dan was sorry for what he had said to Mary and knew he owed her a…'*, the indefinite article *a* is unexpected. In response to the unexpected article there are a few ways that people might recover: either by proposing an alternative noun such as *gift* or *hug* (although the cloze probability of these nouns would presumably be low), or by proposing a consonant-initial adjective in order to maintain their noun original expectation.

The results of norming with the unexpected articles showed that across all sentence contexts, participants used three main strategies to recover from the unexpected articles. For each sentence, then, one of these strategies was used more frequently than the other two, and this will be referred to as the dominant strategy. As we had hypothesized, the most prevalent recovery pattern (the dominant strategy for 70% of the sentences) was indeed for participants to supply an adjective plus the expected noun. For example, for the following sentence,

(1) *Because they were playing baseball so close to the house, the children ended up shattering an…* (where *a window* is the expected completion)

a majority of participants responded with answers such as *(old, expensive, oval, elegant, open, upstairs,* or *antique) window*. In agreement with our "syntactic surprise hypothesis" then, we observed that adjective insertion was the prevalent strategy across most of the sentence contexts.

The next most frequent recovery pattern (the dominant strategy for 20% of the sentences) was for participants to converge upon a single alternate high cloze probability noun. For instance, for the following sentence,

(2) *Whenever Josh had too much to drink he became belligerent and would try to start an…* (where *a fight* would be the expected completion)

participants primarily responded with the noun *argument*.

For the third recovery pattern (the dominant strategy for 10% of the sentences), participants again came up with noun-only responses, except in this case different participants supplied different nouns, i.e. there was no single high cloze probability alternative noun. For the following sentence,

(3) *Lance had moved to Hollywood in hopes of becoming a…*
(where *an actor* would be the expected completion)

participants supplied a variety of noun-only responses such as s*tar, celebrity, superstar, singer,* or *performer*.

From these results, then, it was evident that although adjective insertion turned out to be the predominant strategy, it appeared to be used only when there was not a "good" alternative noun-only response that agreed with the unexpected indefinite article. By "good" I mean that almost all the alternative nouns that participants supplied were either synonyms or highly related nouns, such as *argument* instead of *fight* or *audience* instead of *crowd.* As a side note, the alternative noun-only responses also tended to be within the same category. For instance, for the sentence

(4) *For the snowman's eyes the children used two pieces of coal and for his nose they used an...*

all the noun-only alternative answers were members of the "fruit and vegetable" category (e.g., *artichoke*, *olive*, *asparagus*, etc.). This was an interesting result, considering that they just as easily could have completed the sentence with a word like *icicle* or *acorn*. Instead, they produced words that had overlapping features with the target word. In fact, these results seemed to be compatible with findings from the Federmeier & Kutas (1999) study, which indicated that there was facilitation of words categorically related to constrained sentence targets, as a result of the targets being preactivated.

4.5. Summary

In conclusion, for the norming with unexpected articles, the results indicated that "adjective insertion" at least holds as an off-line strategy, particularly for certain contexts; hence, there was at least the possibility that it might be associated with the noun late positivity observed in Experiments 1 and 2. Whether this is the strategy that individuals actually adopt in the course of on-line processing needs to be tested directly, a challenge we take up in ERP Experiment 3B. Another finding was that the "adjective insertion" strategy seemed to be mediated by the availability of a close semantic associate. Although "adjective insertion" ended up being the prevalent recovery strategy, it seemed to be a backup option. When there was not a high cloze probability noun that fit with the unexpected article, subjects did what they could to preserve their original noun preference. This finding is an intriguing one, though one that is outside the scope of the present research.

EXPERIMENT 3B: ERP STUDY

4.6. Methods

4.6.1. Materials

Using stimuli that exploited a phonological regularity of English indefinite articles – *an's* precede nouns beginning with vowel sounds and *a's* precede nouns beginning with consonant sounds – participants read sentence context pairs of varying constraint that led to expectations for particular vowel or consonant-initial nouns. Stimuli consisted of 160 sentence pair frames with four possible types of target continuation (see **Table 4.2**).

Table 4.2. Composition of experimental and filler stimuli for Experiment 3B.

Condition Number	**Continuation Type**	**Composition**	**Example Context:** *Dale was sorry for what he said to Mary. He knew he owed her…*
1	**E**xpected indefinite **A**rticle-**E**xpected **N**oun (**EA-EN**)	high cloze article + high cloze noun	*an apology*
2	**U**nexpected indefinite **A**rticle-**U**nexpected **N**oun (**UA-UN**)	low cloze article + low cloze noun	*a check*
3	**E**xpected indefinite **A**rticle-**U**nexpected **N**oun (**EA-UN**)	high cloze article + low cloze noun	*an answer*
4 (Filler)	**U**nexpected indefinite **A**rticle-**A**djective-**E**xpected **N**oun (**UA-Adj-EN**)	low cloze article+ + semantically compatible adjective + high cloze noun	*a sincere apology*

The sentence pair contexts were composed in 40 quartets such that each article/noun pair served as a target for each of the three main conditions (EA/EN, UA/UN, EA/UN). Condition 4 continuations (UA/Adj/UN), which were included in order to provide some instances of sentence contexts in which adjectives were actually presented, were always comprised of the same expected noun that was used for the EA-AN condition, preceded by a semantically consistent adjective (primarily determined from norming responses). All indefinite article/noun (adjective) combinations were congruent, i.e. there were no agreement violations between the target articles and nouns or adjectives (i.e., no *a apology* or *an check* target pairs), and the target nouns were all sentence medial. Unexpected noun continuations (EA-UN and UA-UN conditions) were not controlled for plausibility, and ranged from plausible but unlikely to highly implausible.

A total of 640 stimuli were divided into four lists of 160 sentence pairs each, with each participant viewing one list. Sentence pair contexts and article/(adjective)/noun target continuations were used only once per list. Each list contained 40 sentences from each of the four conditions, with equal numbers of *a* and *an* targets.

Approximately one quarter (43) of the items in each list (160) were followed by yes/no comprehension questions pertaining to the immediately preceding sentence.

4.6.2. Cloze procedure

Cloze probabilities were obtained for the 160 experimental sentence pair contexts, norming each context in three different ways: 1) with more probable target indefinite articles provided, 2) with less probable target indefinite articles provided, or 3) truncated prior to the article. Each contextual variation was normed in an off-line sentence completion task by 31-32 University of California, San Diego (UCSD) student volunteers, who were compensated either with experimental credit or cash. In this way, individual cloze probability values were collected for both target articles and nouns. Cloze probability for a given article or noun in a particular context was calculated as the proportion of individuals continuing that particular context with that particular word. Expected indefinite articles had a mean cloze probability of 0.83 (median = 0.87, range = 0.39-1.00), and unexpected indefinite articles had a mean cloze probability of 0.06 (median = 0.00, range 0.00-0.48). Expected nouns alone had a mean cloze probability of 0.88 (median = 0.94, range = 0.28-1.00), UA-UN nouns had a mean cloze of 0.025 (median = 0.00, range = 0.00-0.41), and EA-UN nouns had a mean cloze of 0.00 (median = 0.00, range 0.00-0.16). Adjectives in the UA-Adj-EN condition had a mean cloze probability of 0.06 (median = 0.03, range = 0.00-0.39). Although overall the adjectives exhibited relatively low cloze, they were all semantically congruent within their individual contexts. From our cloze results, we were also able to determine the contexts' constraint values – calculated in the traditional way by using the norming values of the most expected noun for each context. These values will become more relevant as we pursue alternative analyses of our data later in the Results section.

4.6.3. ERP participants

Thirty-two UCSD undergraduates participated in the ERP experiment for course credit or for cash. Participants (23 women, 9 men) were all right-handed, native English speakers with normal or corrected-to-normal vision, ranging in age from 18-23 years, with a mean age of 19.4 years. Ten of the 32 participants reported a left-handed parent or sibling. Three additional participants were excluded from the analysis due to excessive eye blinks or movements.

4.6.4. Experimental procedure

Volunteers were tested in a single experimental session conducted in a soundproof, electrically-shielded chamber. They were seated in a comfortable chair approximately one meter in front of a computer monitor and were instructed to read the stimulus contexts for comprehension. They were also informed that some of the items would be followed by a yes/no comprehension question, to which they were to respond by pressing one of two hand-held buttons. Response hand was counterbalanced across participants and lists. There was a brief practice session that included sentence pairs with both expected and unexpected targets. Participants were asked to remain still during testing and avoid blinking and moving their eyes while the RSVP sentences were being presented. Stimuli were presented in 8 blocks of 20 sentences each. The participants were given a short break after each block.

Sentence stimuli were presented visually in white type on a black background on a cathode-ray tube screen. Each stimulus item was comprised of a context sentence presented in its entirety on screen (with participants instructed to press a button when they had finished reading it) and an RSVP sentence containing the target items. Each centrally presented RSVP sentence began with an empty fixation frame (appearing for a duration that was jittered between 800 and 1300 ms. The fixation frame stayed on-screen over the course of each sentence, with the sentences presented one word at a time in the center of the frame for a duration of 200 ms with a stimulus onset asynchrony of 500 ms. The fixation frame remained on screen between 1.5 and 2.5 seconds following the offset of the sentence final word, after which a comprehension question either did or did not appear in full on the screen. If a comprehension question did appear, the participant's button-press response served to advance the screen to the next sentence. Whether or not there was a comprehension question, there were 2.5 seconds of blank screen prior to the next sentence automatically appearing on screen.

4.6.5. Electroencephalographic recording parameters

The electroencephalogram (EEG) was recorded from 26 electrodes arranged geodesically in an electro-cap, each referenced on-line to the left mastoid. These sites included midline prefrontal (MiPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial frontal (LMFr and RDFr), left and right lateral frontal (LLFr and RLFr), midline central (MiCe), left and right medial central (LMCe and RMCe), left and right mediolateral central (LDCe and RDCe), midline parietal (MiPa), left and right mediolateral parietal (LDPa and RDPa), left and right lateral temporal (LLTe and RLTe), midline occipital (MiOc), left and right medial occipital (LMOc and RMOc), and left and right lateral occipital (LLOc and RLOc). See **Figure 4.4**.

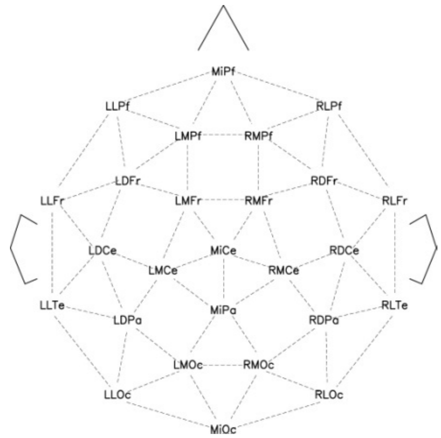


Figure 4.4. Schematic showing the array of the 26 scalp electrodes from which ERPs were recorded.

Blinks and eye movements were monitored from electrodes placed on the outer canthi of each eye and under each eye, each referenced to the left mastoid. Electrode impedances were kept below 5 KΩ. The EEG was amplified with Grass amplifiers set at a band-pass of 0.01 to 100 Hz, and continuously digitized at a sampling rate of 250 Hz.

4.6.6. Data analysis

Trials contaminated by eye movements, excessive muscle activity, or amplifier blocking were rejected off-line before averaging – on average, 7% of data time-locked to the indefinite article and 8% of data time-locked to the noun was rejected. Data with excessive blinks were corrected using a spatial filter algorithm. A digital band-pass filter set from 0.2 to 15 Hz was used on all data to reduce high frequency noise. The data were re-referenced off-line to the algebraic sum of the left and right mastoids and averaged for each experimental condition, time-locked to the onset of the target article and noun.

4.7. Behavioral results

Comprehension accuracy was calculated for the yes/no probe questions. Participants correctly answered an average of 96.1% (median = 97.4%, range = 85% to 100%)

of the questions, indicating that they were attending to and comprehending the experimental sentences during the recording session.

4.8. Cloze probability analysis: ERP results

In the analyses described herein, we take mean voltage measures for the target indefinite articles and the target nouns. For the articles, we analyze the N400 time window (200-500 ms post article onset) in order to attempt to replicate the prediction findings from Experiment 1. For the nouns, several time windows were examined, including: 50-150 ms (frontal N1) and 150-250 ms (P2), in order to probe possible early perceptual/attentional ERP effects of constraint-driven expectancy; 200-500 ms (N400), for investigation of N400 effects and a possible early overlap with the LP effect; 500-1200 ms (an extended LP time window); 500-800 ms (early LP); and 800-1200 ms (late LP). The extended time window for the LP, which might be considered lengthy compared to a more typical 500-800 ms P600 window, was based both on visual inspection of the waveforms from Experiment 1 of this thesis, as well as on time windows used in the more recent "semantic P600" studies. For instance, Nieuwland and Van Berkum (2005) reported using a 700-1300 ms time window to analyze LPs in a recent discourse study, and indeed there seems to be a good deal of latency variability for LPs across "non-syntactic" P600 paradigms. For this additional reason, we also subdivided the extended LP time window into a more typical P600 time window (500-800 ms) and a later time window (800-1200).

The mean amplitude measures in these time windows were subjected to various omnibus analyses of variance (ANOVAs) as well as correlational analyses. All *p*-values provided for ANOVA analyses are reported after epsilon correction (Huynh-Felt) for repeated measures with greater than 1 degree of freedom. Statistical significance ($p \leq .05$)

for correlations (indicated with asterisks) is reported without correction for multiple comparisons.

4.8.1. Target indefinite articles

4.8.1.1. N400 time window (200-500 ms)

To begin, we wanted to confirm whether the prediction ERP effect present at the indefinite articles in Experiment 1 replicated in the current study. An omnibus ANOVA using mean amplitude of the target indefinite articles was conducted during the N400 time window, between 200-500 ms. With 2 levels of article expectancy (Expected articles included those preceding Expected nouns plus those preceding EA-UN nouns; Unexpected articles included those preceding UA-UN nouns as well as those preceding adjectives in the filler condition) X 26 levels of electrode, there was a main effect in the anticipated direction [$F(1,31) = 4.52$, $p = .0416$], with Unexpected articles showing increased negativity relative to Expected articles (a .38 μV difference). See **Figure 4.5a, b** and **c**.

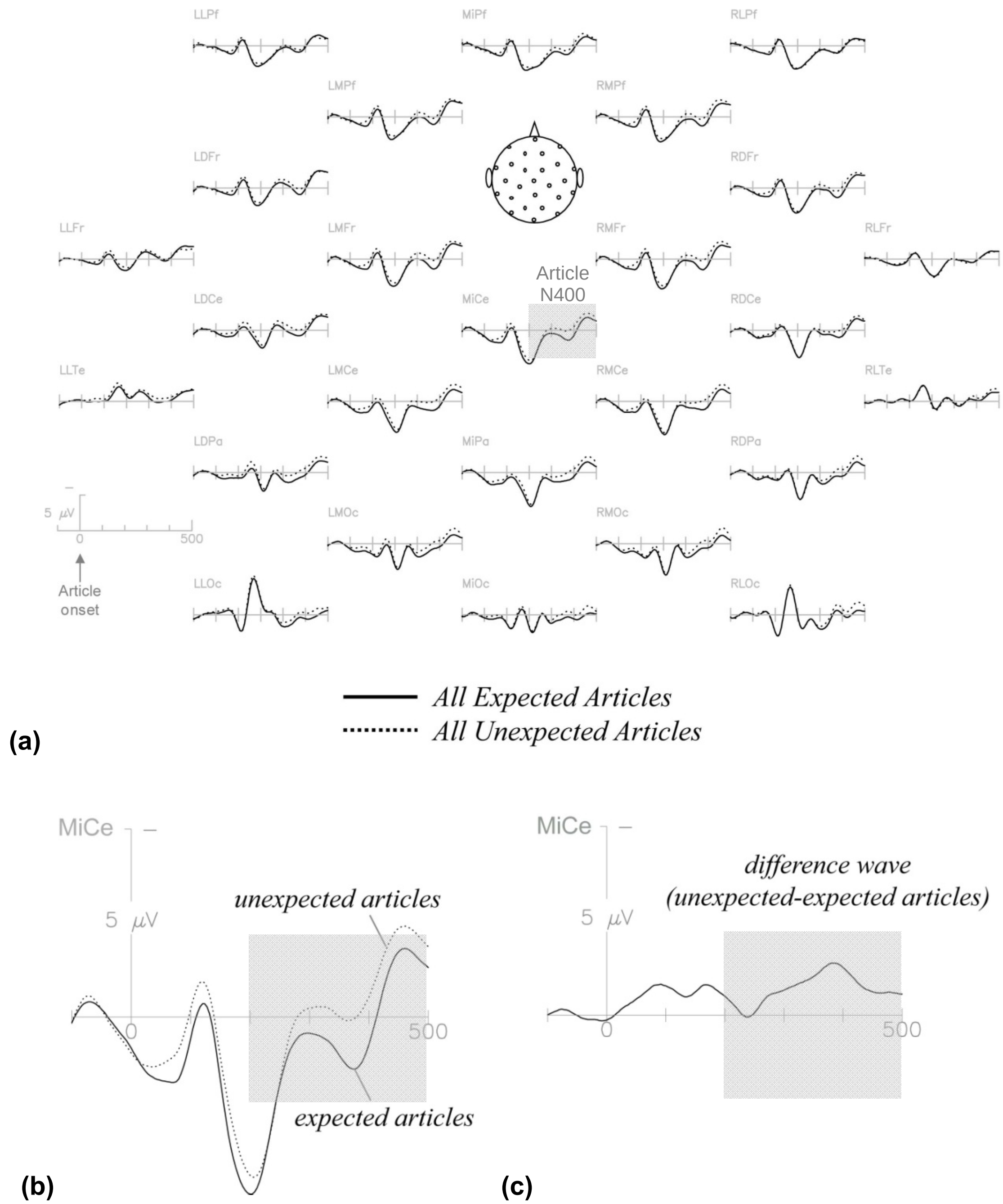


Figure 4.5. ERPs to more and less expected target indefinite articles, with article N400 time window highlighted: (a) over all 26 scalp electrodes, (b) at a single electrode MiPa, and (c) shown as a difference wave (less minus more expected) at the vertex electrode, MiCe. In (a) and (b), more expected articles are indicated with solid line, less expected with dotted line.

In addition to ANOVAs, we conducted a correlation analysis using conditions identical to that of Experiment 1 in which article cloze probability was correlated with article mean amplitude. For each item, articles were sorted according to their cloze probability into ten equal-width bins spanning the interval 0-100%. In each such bin, the mean cloze probability and mean ERP amplitude in the N400 window (200-500 ms) were computed. Correlation values over the 26 scalp electrode sites are show in **Figure 4.6**. Correlation patterns are similar to those observed in the original study, albeit with slightly lower overall *r*-values. However, the overall correlations resemble those of Experiment 1 in that as cloze probability decreases, mean amplitude in the N400 time becomes more negative, with maximal *r*-values clustering at electrode sites where the N400 effect is usually largest (posterior scalp areas).

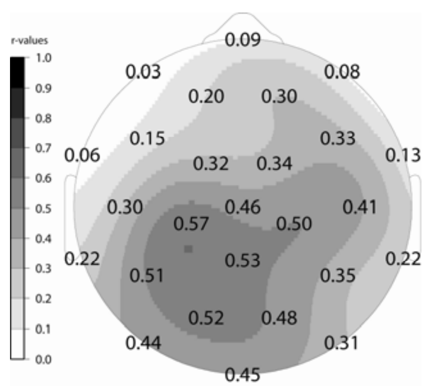


Figure 4.6. Correlations (*r*-values) of indefinite article mean amplitude (200-500 ms) with offline article cloze probability. No statistically significant (p ≤ .05) correlations were present in this analysis.

4.8.2. Target nouns (and adjectives)

Figure 4.7 plots a time window (-100 to 700 ms) encompassing the frontal N1, frontal P2 and N400 ERP effects for the Noun and Adjective conditions. For the N1 and P2 analyses, ERP mean amplitudes of the three main experimental conditions plus the Adjectives were contrasted: Expected, UA-UN, EA-UN, Adjectives (see **Figure 4.8**). The Adjective condition was included in these analyses with the assumption that these early components might reflect to a greater degree the influence of the Unexpected or Expected indefinite articles that preceded the target nouns or adjectives, rather than the word class features/cloze probability of the target stimuli themselves. The Adjective condition was also included in some of the N400 analyses to ensure that N400 amplitude of the (relatively) low cloze adjectives was generally being modulated in the expected way, though no principled predictions were formed for this condition. For the LP analyses, the Adjective condition was not included for a few reasons. First, the fact that the words following adjectives are target nouns, while in the other three conditions words following target nouns are for the most part closed-class words, makes them difficult to compare. Second, in the case of the Adjective condition, resolution of an expectancy violation would be delayed by 500 ms, due to the additional word (article+adjective+noun compared to simply article+noun), also making potential ERP differences difficult to interpret. **Figure 4.9** plots an extended time window (out to 1500 ms post noun onset) for all four experimental conditions.

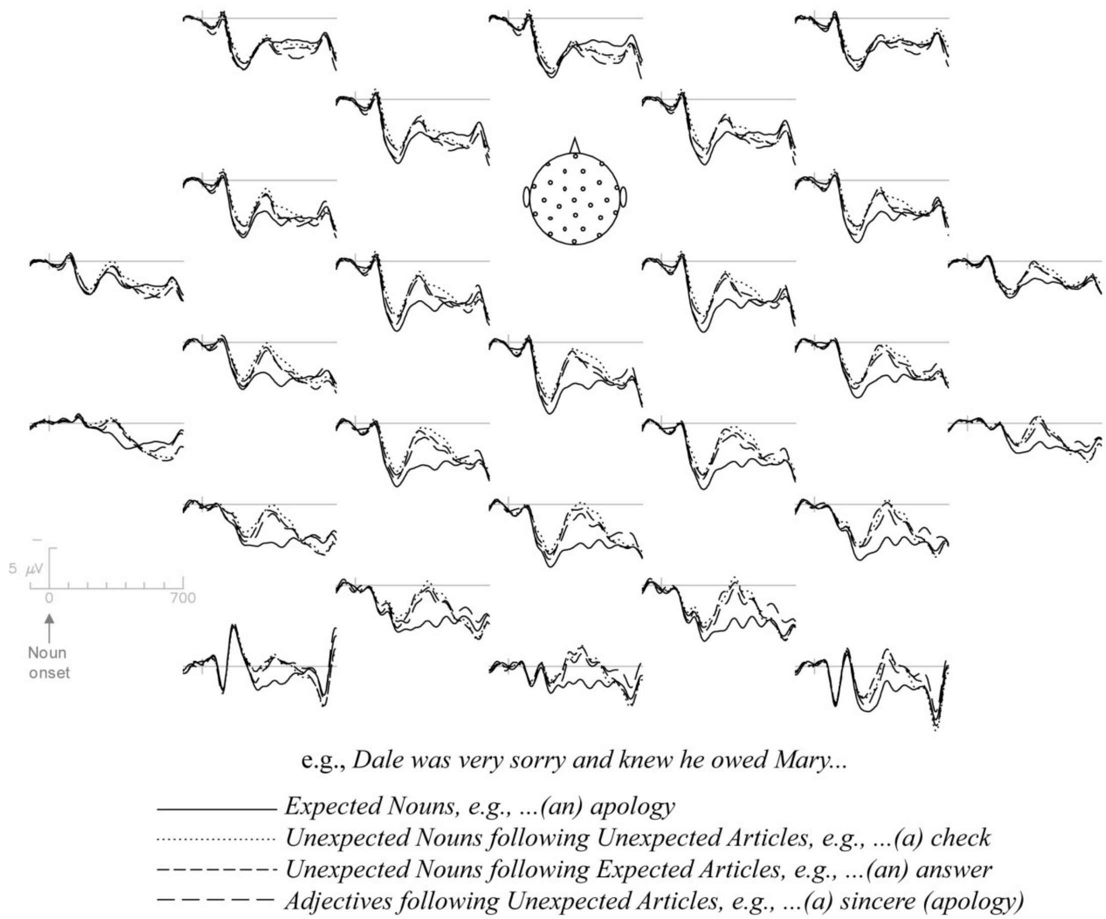


Figure 4.7. ERPs for cloze probability analysis, time-locking to nouns and adjectives: -100-700 ms time window encompassing frontal N1 (50-150 ms), frontal P2 (150-250 ms), and N400 (200-500 ms) effects over all 26 electrode sites.

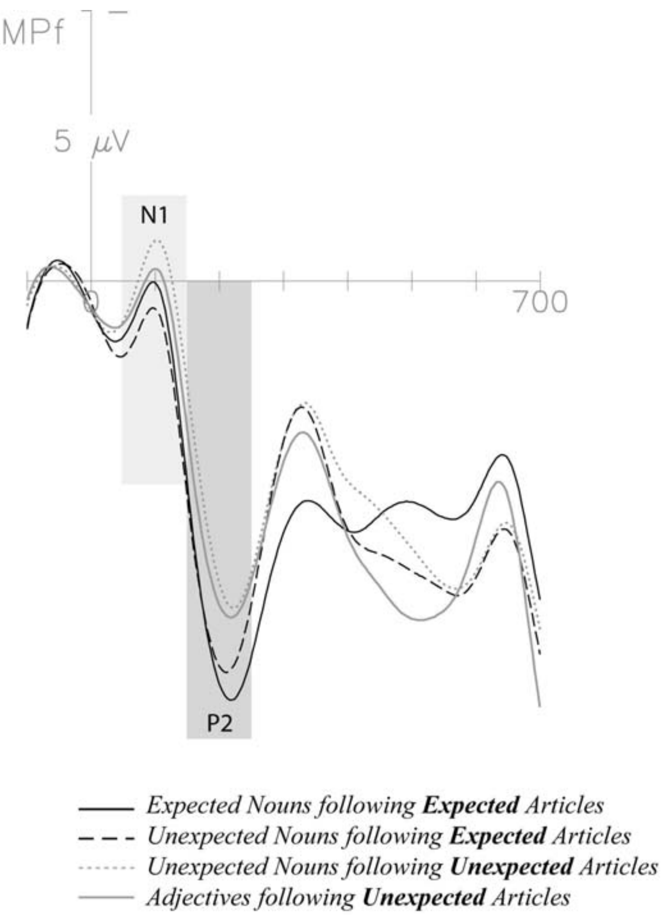


Figure 4.8. Cloze probability analysis, time-locking to nouns and adjectives: time window encompassing N1 and P2 effects at a single frontal electrode site (LMPf).

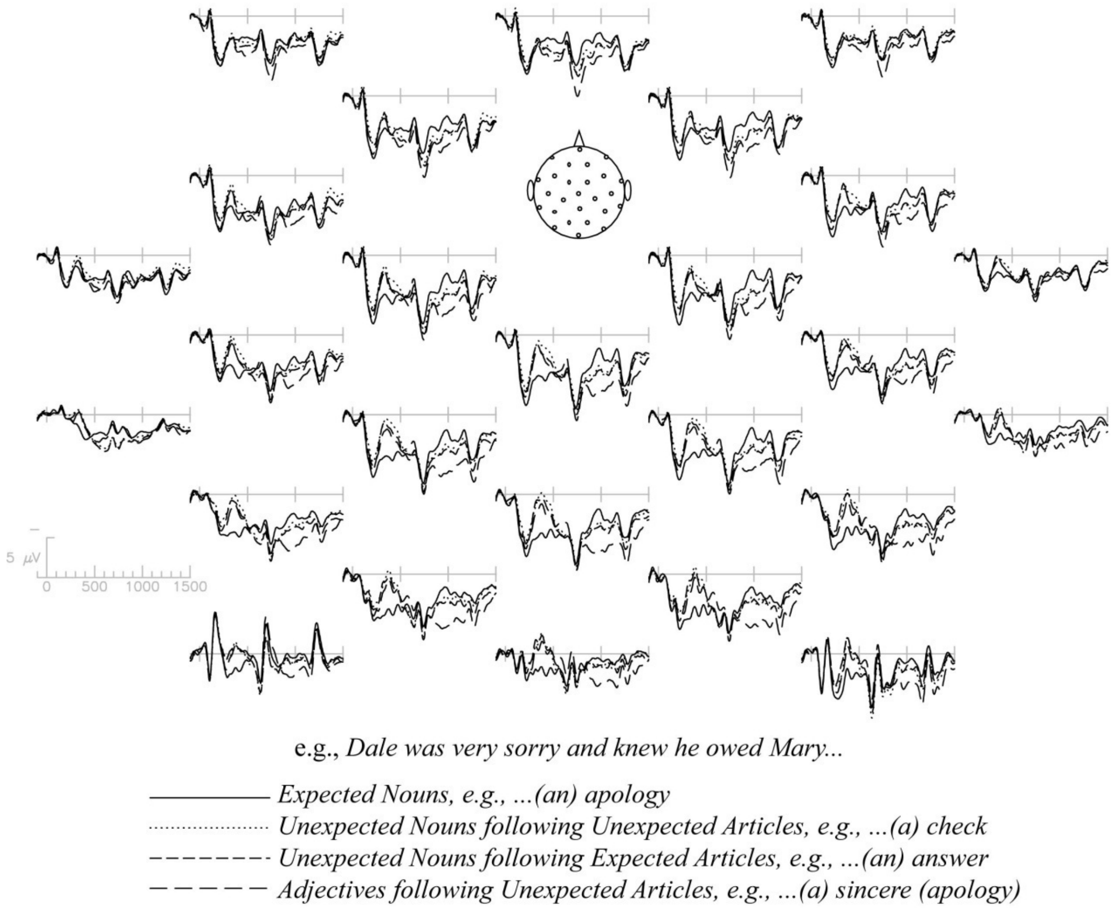


Figure 4.9. Cloze probability analysis, time-locking to nouns and adjectives: -100 to 1500 ms time window over all 26 electrode sites.

4.8.2.1. Frontal N1 (50-150 ms)

To examine possible influences of Noun/preceding article type on the frontal N1, we took voltage measures from 11 frontal scalp electrodes (MiPf, LLPf, RLPf, LMPf, RMPf, LDFr, RDFr, LLFr, RLFr, LMFr, and RMFr). **Figure 4.10** indicates these channel locations.

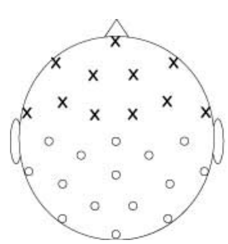
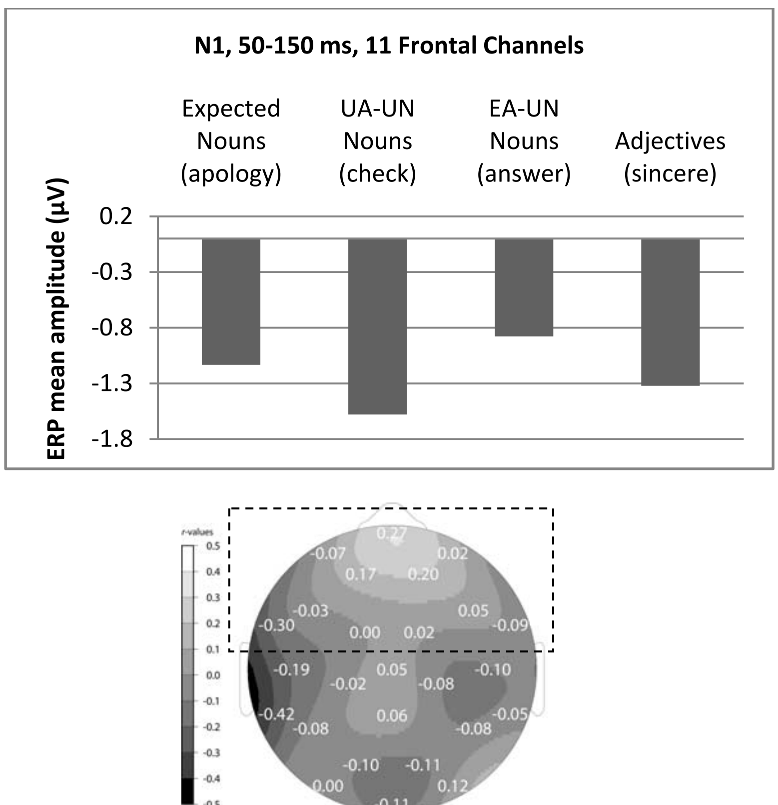


Figure 4.10. Locations of the 11 frontal electrode sites used for N1 and P2 analyses.

These measures were subjected to an omnibus 4 (target type: Expected Noun, UA-UN, EA-UN, Adjectives) X 11 (electrode site) ANOVA. There was a main effect of target type [$F(3,93) = 2.94$, $pHF = .0377$], with UA-UN nouns and Adjectives showing greater negativity (both preceded by Unexpected articles) compared to Expected and EA-UN nouns (both preceded by Expected articles), see **Figure 4.11a.** Bonferroni-corrected pairwise comparisons revealed that the mean amplitude of UA-UN nouns differed significantly from that of the two noun types preceded by expected articles ($p<.01$ with Expected nouns, $p<.0001$ for EA-UN nouns), though Adjectives differed significantly only from the EA-UN nouns ($p<.01$). This enhanced N1 to items following Unexpected articles may be indicative of an enhanced attentional load when the articles signal a surprising continuation.



(a) (b)

Figure 4.11. N1 (50-150 ms) time window: (a) ERP mean amplitude of noun and adjective conditions, and (b) correlations of N1 amplitude with noun cloze probability, with frontal electrode sites included in statistical analyses highlighted in box. No statistically significant (p ≤ .05) correlations were present in this analysis.

To determine if there was any influence of noun cloze probability on the frontal N1, noun cloze was correlated with noun mean amplitude (50-150 ms) for the Expected and Unexpected nouns. As **Figure 4.11b** reveals, the correlations were weak over nearly all 11 electrode sites where the ANOVA analysis was conducted. Given the results of the ANOVA, where N1 amplitude patterned with the expectedness of the preceding article regardless of the contextual fit of the target noun/adjective, these results are not so surprising.

4.8.2.2. Frontal P2 (150-250 ms)

To investigate the potential influence of cloze probability/preceding article type on noun frontal P2 amplitude, we analyzed ERP data from the same 11 frontal electrodes and same four levels of Target Type as for the N1 analysis. Frontal P2 responses have been linked to higher order perceptual and attentional processes, and enhanced frontal P2s have been observed to semantic constraint/cloze manipulations. There was a main effect of Target Type [$F(3,93) = 8.99$, $pHF = <.0001$] with Expected and EA-UN nouns showing enhanced P2s (both preceded by Expected articles) compared to UA-UN nouns and Adjectives (both preceded by Unexpected articles), see **Figure 4.12a**. Bonferroni-corrected pairwise comparisons revealed that two Target Types preceded by Expected articles (Expected and EA-UN nouns) each differed significantly (all comparisons $p<.001$) from the two Target Types preceded by Unexpected articles (UA-UN nouns and Adjectives). Target Types preceded by similar article types did not differ significantly from each other.

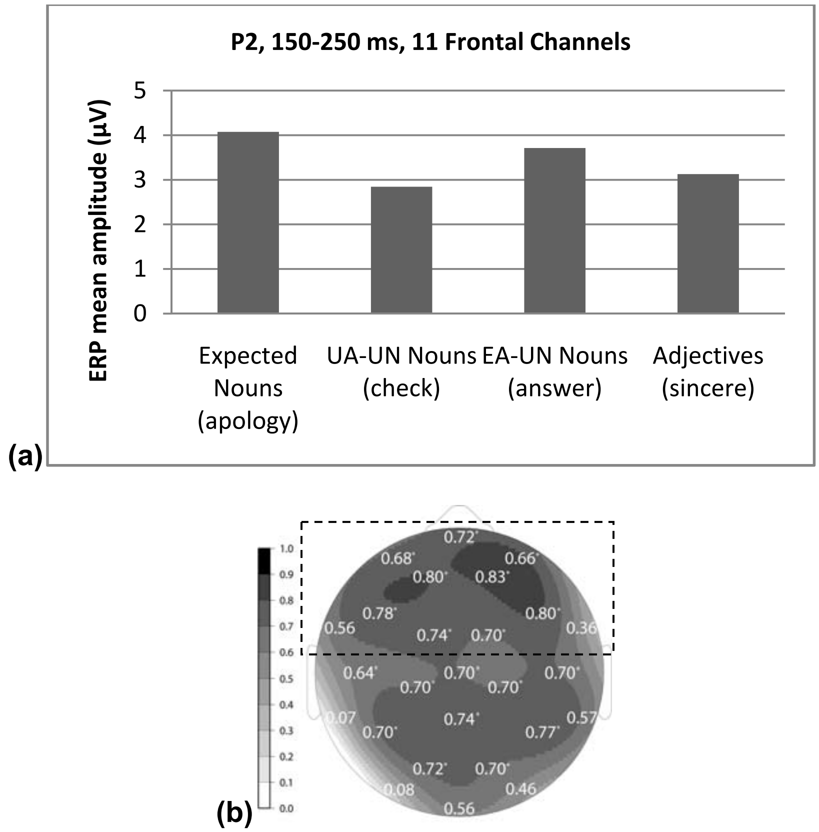


Figure 4.12. P2 (150-250 ms) time window (a) ERP mean amplitude of noun and adjective conditions, (b) correlations of P2 amplitude with noun cloze probability, with frontal electrode sites included in statistical analyses highlighted in box. Significant correlations (p≤.05) are indicated with asterisks (*).

We also performed a correlation analysis of cloze probability with the mean amplitude for the P2 of the Expected and Unexpected nouns, similar to that done for the N1, except between 150-250 ms. Correlations at individual electrode sites are plotted in **Figure 4.12b**. There was a pattern of strong positive *r*-values present at 11 frontal sites (maximal *r*-value was 0.83 at RMPf) where the P2 ANOVA analysis was conducted. These positive *r*-values indicate that there is an increase in P2 mean amplitude in this time window with increasing cloze probability. At first, these results may seem at odds with the ANOVA results for the same time window, where P2 mean amplitude seems to reflect the buildup of contextual constraint as a function of the preceding article, rather the cloze of the actually presented noun. However, constraint and cloze probability are confounded on one end of the spectrum (high constraint is defined by a single high cloze continuation), though not at the other (low cloze items can continue either high or low constraint contexts). In combination, these results suggest that an analysis in which the low cloze items are sorted according to their contextual constraint may allow effects of cloze and constraint to be dissociated, in order to determine which factor influences P2 amplitude to a greater degree.

4.8.2.3. N400 time window (200-500 ms)

In order to determine if the noun N400 findings from Experiment 1 were replicated and to ensure that the Expected versus Unexpected nouns (as well as the Adjectives) exhibited the anticipated, typical N400 waveform patterns, an omnibus ANOVA was conducted using mean amplitude of the Expected noun, the two Unexpected (UA-UN, EA-UN) noun and the Adjective conditions in the 200-500 ms time window. See **Figure 4.13**. There was a main effect of Noun/Adjective Type [$F(3,93) = 27.85$, $pHF = <.0001$], with the mean amplitude of UA-UN nouns showing the greatest negativity (0.71 μV), followed by EA-UN nouns (1.03 μV), followed by Adjectives (1.50 μV) compared to Expected nouns (2.72 μV). Bonferroni-corrected pairwise comparisons of all four levels revealed significant differences between each level ($p<.0083$). These results of the Expected versus UA-UN nouns, in particular, replicate the earlier findings. The fact that the two Unexpected nouns and the Adjectives differ significantly from each other will be addressed in the discussion section.

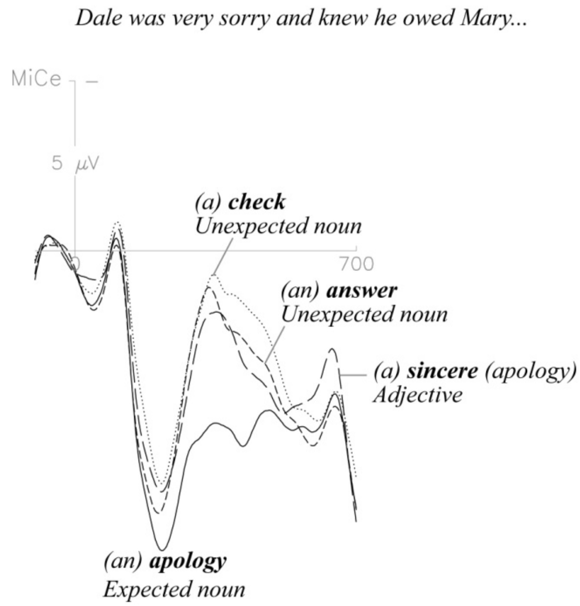


Figure 4.13. N400s to Expected and Unexpected target nouns at the vertex electrode (MiCe).

Correlation analyses for the Expected and UA-UN nouns were also conducted (using the conditions identical to those of Experiment 1) in which noun cloze probability was correlated with article mean amplitude between 200-500 ms post-noun onset. The *r*-values over the 26 scalp electrode sites are shown in Figure 4.14. Correlation patterns are similar to those observed in the original study: when noun cloze probability decreases, noun mean amplitude in the N400 time window decreases, with widely distributed high correlation values, but with maximal *r*-values clustering over typical N400 sites (i.e., centro-parietal scalp locations).

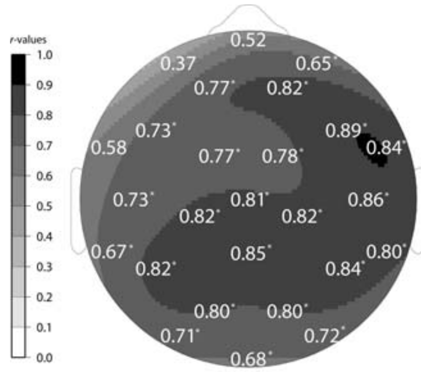


Figure 4.14. Target noun correlations (*r*-values) of noun mean amplitude (200-500 ms) with noun cloze probability, replicating N400 correlation patterns from Experiment 1. Significant correlations (p ≤ .05) are indicated with asterisks (*).

4.8.2.4. LP time windows (500-1200 ms, extended LP; 500-800 ms, early LP; and 800-1200 ms, late LP)

An ANOVA (with 26 levels of Electrode) on the mean amplitude of the Expected and two Unexpected noun types in the extended LP time window (500-1200 ms) revealed a main effect of Noun Type [$F(2,62) = 6.17$, $pHF = .0038$], with the two Unexpected conditions (UA-UN=2.46μV, EA-UN=2.60μV) showing an enhanced LP relative to the mean amplitude of Expected nouns (1.73μV). See **Figures 4.15a, b, and c**.

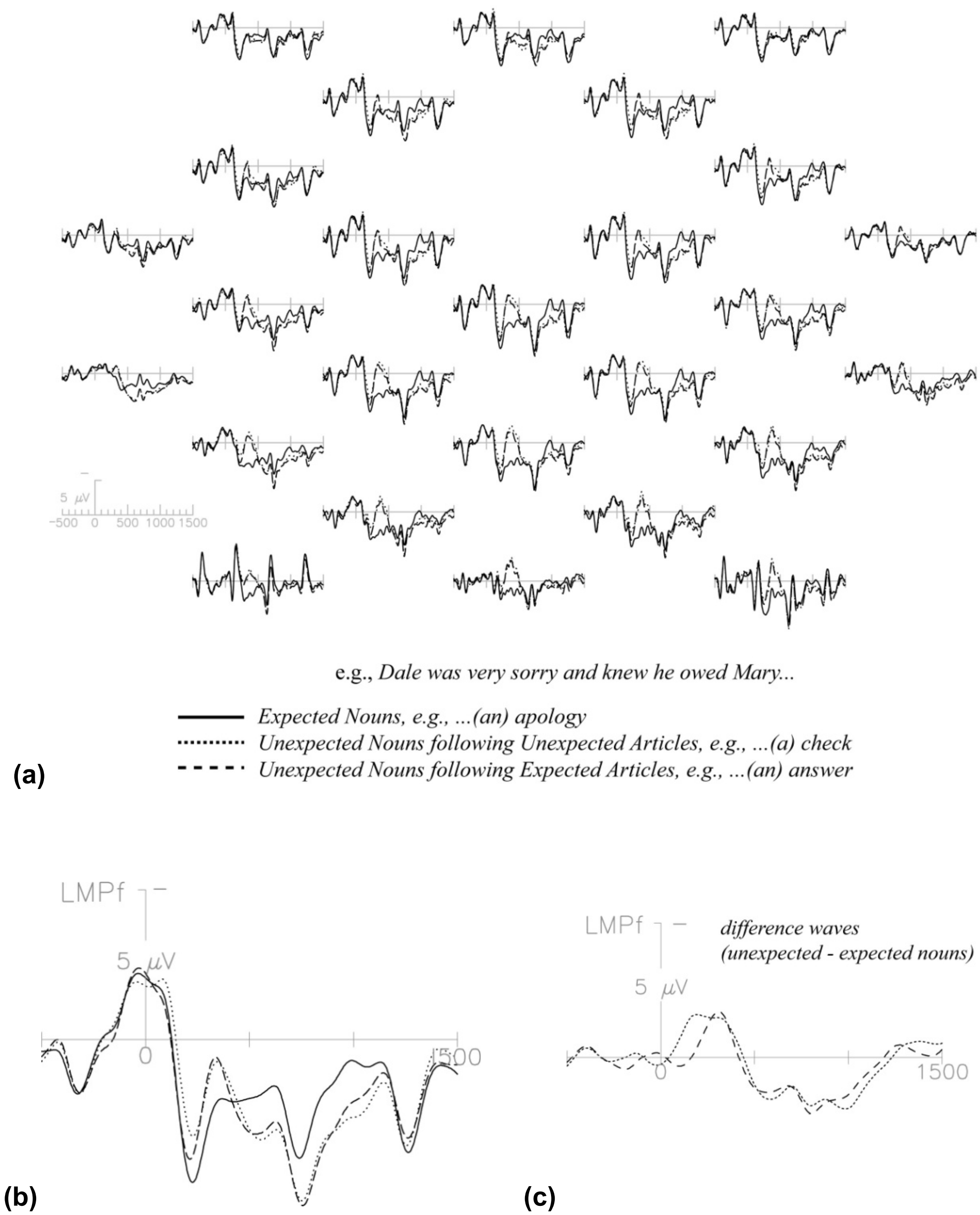


Figure 4.15. Epoch including LP time window (500-1200 ms) to Expected and Unexpected target nouns (a) at all 26 scalp electrodes, (b) at a single electrode LMPf, and (c) difference waves (less-more expected) at LMPf. For the difference waves, (UA-UN minus Expected) is indicated with a dotted line, (EA-UN minus Expected) with a dashed line.

Planned pairwise comparisons of the three levels revealed significant differences between the Expected and both of the Unexpected noun conditions (each with $p<.01$), with the two Unexpected noun types not differing significantly (p = 0.11). This same pattern of significant differences between both Unexpected noun types and the Expected nouns but not between the two Unexpected noun types also occurred in both the early (500-800 ms) and late (800-1200 ms) LP time windows. Both Unexpected noun types (*check* and *answer*) elicited larger LPs relative to the Expected nouns over the 26 electrode sites. To explore the distributional nature of this effect, for the extended LP (500-1200 ms) separate ANOVAs were conducted for each Unexpected Noun Type contrasted with Expected nouns (2 levels of expectancy X 26 levels of electrode site). In both cases, there were significant interactions of expectancy with Electrode site, which were followed up via distributional analyses using a subset of 16 representative channels across the scalp. These locations were divided into: 2 levels of hemisphere (left vs. right) X 2 levels of laterality (lateral vs. medial) X 4 levels of anteriority/posteriority (prefrontal vs. frontal vs. parietal vs. occipital). **Figure 4.16** indicates these electrode locations.

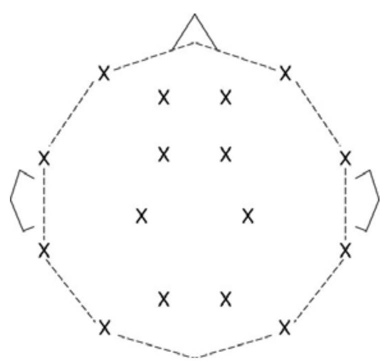


Figure 4.16. Locations of the 16 electrode sites used for the distributional analyses.

For both analyses there were significant interactions of expectancy X laterality, though both were mediated by a higher-level interaction with anteriority: medially the LP effect was progressively larger from posterior to anterior sites (largest at Prefrontal and Frontal), but largest at Central sites laterally. Overall, then, the findings and the scalp distributions of the LP to UA-UN nouns like *check*, replicated our (unanticipated) LP findings from Experiment 1. The condition introduced in the current study, the EA-UN nouns (*answer*), also exhibited a similar LP waveform pattern to that found for the UA-UN nouns (*check*) in the previous, as well as current, study, and disconfirmed our central hypothesis that a differential LP pattern for this condition would be indicative of "syntactic surprise".

In addition to the ANOVA analysis on the LP time window, we also wanted to verify that noun LP mean amplitude was *not* correlated with noun cloze probability. While it was not anticipated that the LP amplitude, like N400 amplitude, would vary systematically with cloze, the novelty of our LP finding warranted testing for this possibility. **Figures 4.17a, b,** and **c** show *r*-values for such correlations over the LP time windows. If LP mean amplitude and noun cloze probability are correlated, then we would expect a pattern of significant negative correlations, where as cloze decreases, LP amplitude increases. However, in all three LP time windows the correlation strengths at individual electrodes were generally weak, with only a few isolated electrode sites where *r*-values do approach significance (e.g., *r* = -.55 at the MiCe site in the 800-1200 ms time window). Furthermore, the scalp distributions of the maximal correlations did not map onto the topographical areas where the LP effect was observed in the distributional analysis described above (medial, prefrontal and frontal sites).

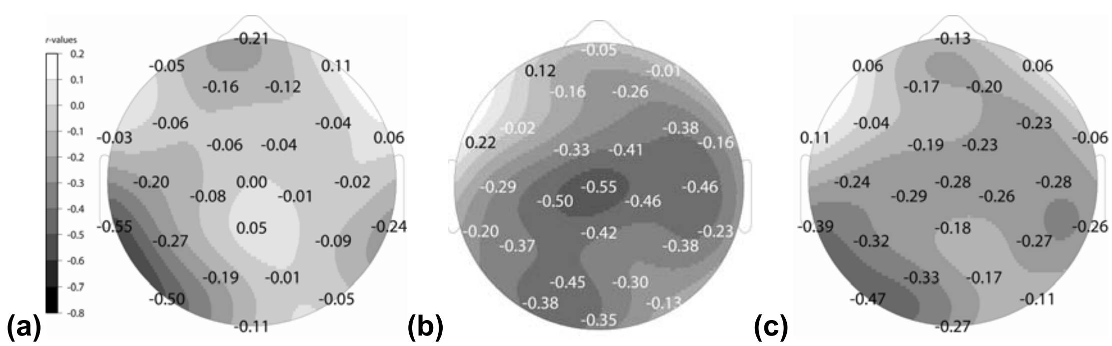


Figure 4.17. Noun cloze correlated with LP mean amplitude (a) 500-800 ms (b) 800-1200 ms (c) and 500-1200 ms. No statistically significant (p ≤ .05) correlations were present in these analyses.

4.9. Cloze probability analysis: discussion

We will now provide a brief summary of the experimental results for the cloze probability analyses performed up until this point: (1) replication of indefinite article and noun N400 effects from Experiment 1, (2) the early ERP effects (frontal N1 and P2) for the target nouns/adjectives, (3) new N400 findings at the target nouns/adjectives, and (4) a preliminary discussion of the effects observed in the LP time window.

4.9.1. Replication of previous N400 effects

To summarize the N400 effects, replicating the results of our original study at both the articles and nouns in the N400 time window confirmed that a) probabilistic prediction of specific word forms was again evidenced by the graded N400 effect at the article, and b) the pre-conditions for eliciting the noun LP as observed in Experiment 1 were in place. With a range of off-line expectancy for both the articles and nouns, it was clear that the N400 was indexing contextual fit in the usual way.

4.9.2. Frontal N1 and P2 for noun/adjective targets

The ERP results for the target nouns (and adjectives) between 50-150 ms and 150-250 ms post onset indicate that there appears to be reduced N1 and enhanced P2 amplitude for targets following Expected relative to Unexpected indefinite articles. Based on findings of enhanced frontal N1s to targets in attended locations (in spatial attention paradigms) and findings of reduced N1s to expected sentence continuations in high versus low constraint contexts (in sentence ERP studies), our results seem to accord with the idea that the frontal N1 reflects allocation of attentional resources. In the present study, target indefinite articles that signal an unexpected continuation (and therefore contribute to lower sentential constraint) may serve to "cue" the attentional system in preparation for what is about to follow. As for the frontal P2 component (often associated with aspects of high-order visual processing), amplitude increases have previously been noted in semantic language ERP experiments for highly constrained relative to weakly constrained targets (e.g. Federmeier & Kutas, 2001; Federmeier, Mai & Kutas, 2005; Wlotko & Federmeier, 2007). Our finding of larger amplitude frontal P2s for word targets following expected articles (i.e., high constraint contexts) relative to unexpected articles patterns similarly with these previous studies: regardless of the cloze probability of the target noun or adjective, P2 amplitude was apparently influenced by constraint (inferred from the indefinite article, because in this analysis, constraint was not measured directly). Prediction-consistent indefinite articles seemed to guide the degree to which processing resources were devoted to visual feature extraction, such that when an upcoming target noun was expected, the system was primed to receive visual information of a certain form.

4.9.3. New N400 findings: target indefinite articles and nouns

One unanticipated finding was that Unexpected nouns preceded by Unexpected articles (UA-UN nouns) exhibited larger N400s than Unexpected nouns preceded by Expected articles (EA-UN nouns) and both showed significantly larger N400s than the Adjective condition. One explanation is that the enhanced negativity in the UA-UN *article* N400 time window may have continued into and influenced the noun time window. Additionally, the smaller N400s to Adjectives relative to both UA-UN and EA-UN nouns may simply be a function of their slightly higher cloze probabilities. We do not believe that this finding alters the main conclusions we would like to make about the present data.

4.9.4. Late positivity time window: target nouns

Regarding the LP, we wanted to determine whether the effect would be present to one or both Unexpected noun conditions. We had originally proposed that an enhanced LP relative to Expected nouns in the UA-UN but not the EA-UN nouns would be compatible with the idea of the LP indexing a syntactic expectancy violation. The finding that the two Unexpected noun conditions exhibited similar prolonged LPs following N400s relative to

Expected nouns disconfirmed this hypothesis. These results suggested that rather than the LP reflecting a violation of a contextual expectancy for an upcoming part of speech, which the parser might activate once an unexpected indefinite article appears, the positivity instead seems to be reflecting a similar process in *both* unexpected conditions. In other words, the LP is present to the Unexpected nouns *regardless* of whether such nouns are preceded by prediction-consistent or -inconsistent articles.

4.10. Intermediate conclusions and proposals

Our essentially null finding of similar amplitude and distribution LPs to both Unexpected noun types (along with early N1 and P2 effects consistent with the idea that contextual constraint is a factor shaping the comprehension process) argued for an alternative interpretation. Our results suggested that the LP may index some kind of (non-syntactocentric) processing “cost” when the context highly constrains upcoming input but an unexpected lexical item is received that violates that prediction. In other words, the LP originally observed in Experiment 1 to contextually unexpected noun targets might be related to a more general “failure” of prediction and represent processing difficulty when the parser realizes it has predictively headed down the wrong path. Under this alternative hypothesis, an LP difference between the two unexpected noun continuation types would *not* be expected, which is in line with our actual results. This explanation would also be more compatible with recent findings of late (500-900 ms) frontal positivities following N400s by Federmeier, Wlotko, De Ochoa-Dewald & Kutas (2007) to unexpected endings in strongly, but not weakly, constraining sentence contexts – in other words, to a strictly semantic manipulation. In terms of the type of probabilistic prediction outlined in Experiment 1, one could extrapolate that such “costs” might arise when an item has been strongly activated in advance of the input, with extensive support from the available sources of constraint (e.g., lexical associations, on-line sentence or discourse context construction, schema, event structure, world knowledge, syntax, or phonological matching, to name a few), but those items are not received. At this point we should clarify what we mean – or at least what we do not mean – by terms such as “prediction cost”, “misprediction” or even “constraint violation”. These terms might seem to imply a sort of conscious processing that highlights the negation of a single, highly-activated item, rather than a range of variably activated, but possible, continuations. But even in sentence contexts considered moderately or not very constraining (e.g., *There was nothing wrong with the…car*), some continuations will be more likely than others: for instance *car* in this isolated context is a more probable completion than *carburetor, orange, excitement, greedy*, or *about*. (Note that a gradation in fit of alternative continuations is difficult to quantify, because in the standard, but inadequate, assessment of contextual expectancy – the offline cloze probability norming task – none of these items is likely to be provided, and thus they would all share the same low or zero percent cloze probability rating.) So if there is a “cost” to receiving an unexpected item when a different item(s) has been pre-activated, one might expect this effect to be most evident to unexpected continuations in highly constraining contexts, where there is a single, highly probable continuation (e.g., *He mailed the letter without a…thought* where *stamp* is expected). Additionally, we project that such costs might not only be evidenced when constraint is maximal, but that there might be cost even when constraint is not so high, for instance when there are a number of possible continuations which have been preactivated, though none to a strong degree. However, the question of whether “cost” differences would be elicited by unexpected continuations in both strongly and not-as-strongly constraining contexts is an empirical one.

Based on extensive N400 literature dating back as far as Kutas & Hillyard’s 1984 study, we know that the N400 does *not* exhibit a sensitivity to the degree of constraint violation – and to reiterate, by violation, here, we mean the degree of mismatch between a sentence's constraint value and the cloze of the actually presented item. Instead, the N400 is sensitive to the cloze probability of a presented item in a particular context. Historically, the lack of experimental evidence for a consequence to mispredicting has made anticipatory language processing an unappealing proposal among many psycholinguists. With a lack of behavioral evidence for a "cost" to processing unexpected sentence continuations, and the N400's insensitivity to "misprediction", a general question of interest is whether there is *any* evidence for constraint violation? To date, a strong case has not been made, but with respect to the current study's results, we wonder whether our LP may in fact reflect such a cost. If there is modulation of LP amplitude to low cloze probability sentence continuations as a function of how constraining the eliciting context is, then this would serve as strong evidence for a "cost" for (mis)prediction. Furthermore, finding two-pronged evidence for prediction and its costs (framed here in terms of a double dissociation between the sensitivities of the N400 and the LP) would certainly help to substantiate this point.

The above considerations motivated us to explore the possibilities that: (1) the LP effect indexes violation of sentential constraint, with constraint being operationalized as the cloze probability of the most contextually expected noun (as determined through stimulus norming), (2) similar to our graded prediction effect, there may be a gradient of cost depending on the degree of constraint violation, and (3) there could be a double dissociation between the sensitivities of the N400 (to cloze probability, i.e., fit within a context) and the LP (potentially to the degree of constraint violation). To investigate these ideas, we performed additional analyses on the ERPs in the LP time window, using more fine-grained breakdowns of the stimuli, sorting on various measures of constraint. Specifically, we wanted to observe whether LP amplitude increased for Unexpected sentence continuations as a function of how constraining the preceding context was.

Though our experiment was not explicitly designed to test for effects of contextual constraint, and therefore did not manipulate this factor in a systematic way, we took advantage of the range of constraint afforded to us by the stimuli. Within our available data, noting a pattern of results in which LP amplitude is modulated by degree of constraint violation would be compatible with a predictive view of language processing. A model of the LP as an index of constraint violation would also present a challenge to the current views of the P600 component (assuming a degree of similarity between our LP and the "canonical" P600), both in light of its traditional interpretation as an index of syntactic processing, as well as its more recent explanations in terms of reflecting various semantic-related processes.

4.11. Sentential contextual constraint analyses: ERP results

4.11.1. Traditional contextual constraint analysis

To explore how the LP might relate to violation of sentence constraint, we examined ERP effects to Unexpected nouns as a function of how constraining their preceding contexts were. Because constraint and cloze probability are confounded in the upper range of cloze probability (i.e., by definition a high cloze probability continuation defines a highly constraining context), we limited ourselves to examining only the Unexpected noun continuations, where preceding sentence constraint could vary, but (low) cloze probability of the continuation is held relatively constant. We began by using a traditional measure of constraint; that is, we determined contextual constraint by calculating the cloze probability of the most commonly provided norming responses when the expected articles were provided (**Figure 4.18**).

"Traditional" contextual constraint measures. Determined by norming results with Expected articles:

e.g., *Dale was very sorry and knew he owed Mary an...*

Expected Noun determines constraint:
apology = 97% cloze probability, constraint=97%

Unexpected Nouns: UA-UN: *(a)* ***check*** = <1% cloze probability
EA-UN: ***answer*** = <1% cloze probability

Figure 4.18. Traditional constraint measure, sample stimulus.

Sentence stimuli ranged in the cloze probability of the most expected nouns (i.e., had different levels of sentence constraint). For instance:

(5a) High constraint context/High cloze continuation

Dale was sorry for what he said to Mary. He knew he owed her an...

...apology

(97% cloze of most common completion *apology* determines constraint= 97%)

(5b) High constraint context /Low cloze continuation

Dale was sorry for what he said to Mary. He knew he owed her an...

...answer or *(a) check*

(97% constraint, both Unexpected nouns = 0% cloze)

(6a) Lower constraint context /Medium cloze continuation

Bo began his military service as an enlisted soldier. After completing his service academy training he became an...

...officer

(58% cloze of most common completion *officer* determines constraint = 58%)

(6b) Lower constraint context /Low cloze continuation

Bo began his military service as an enlisted soldier. After completing his service academy training he became an...

... artist or *(a) professor*

(58% constraint, both Unexpected nouns = 0% cloze)

We used these values to correlate constraint with mean amplitude of the Unexpected nouns for both of the Unexpected noun types, UA-UN and EA-UN, since both of these conditions showed enhanced LP following the N400 relative to the Expected nouns. Because the range of sentence constraint in this experiment was limited (i.e., all but 3 of the 160 sentence contexts had constraint values greater than 50%, such that there were no very low constraint contexts) we were unable to explore Unexpected endings over a full (0-100%) range of constraint. Based on the available contextual constraint values, we sorted the sentence stimuli into 5 10% bins spanning the interval 50-100% and correlated the ERP mean amplitude of the bins between 500-1200 ms post-noun onset with the average constraint value per bin, separately for UA-UN and EA-UN nouns. We were assured that any significant correlation patterns would not be due to the cloze probabilities of the actually-presented Unexpected nouns, because the cloze values of the different bins (on average) were all similarly low (see **Table 4.3**).

Table 4.3. Traditional constraint analysis breakdown of constraint conditions.

Constraint Bin	Number of Items per Bin	Average Bin Constraint	Average cloze of presented UA-UN nouns	Average cloze of presented EA-UN nouns
90-100%	95	.96	.03	.00
80-89%	39	.84	.01	.00
70-79%	13	.76	.04	.01
60-69%	6	.68	.00	.00
50-59%	4	.56	.00	.00

4.11.1.1. Correlations of traditional contextual constraint measure with ERP mean amplitude of Unexpected nouns over various time windows

Though our primary predictions centered on possible findings in the LP mean time window(s), we performed correlations of contextual constraint with UA-UN and EA-UN mean amplitude over six different time windows: N1 (50-150 ms), P2 (150-250 ms), N4 (200-500 ms), early LP (500-800 ms), late LP (800-1200 ms), and extended LP (500-1200 ms). See **Figures 4.19a-g and 4.20a-g**. We tracked the progression of the constraint violation effect in this way in order to better isolate the timing, including the onset, of the ERP's sensitivity to this factor. We additionally wanted to follow up on some of the earlier findings from our cloze analyses which suggested that ERP patterns in the early time windows may be related contextual constraint.

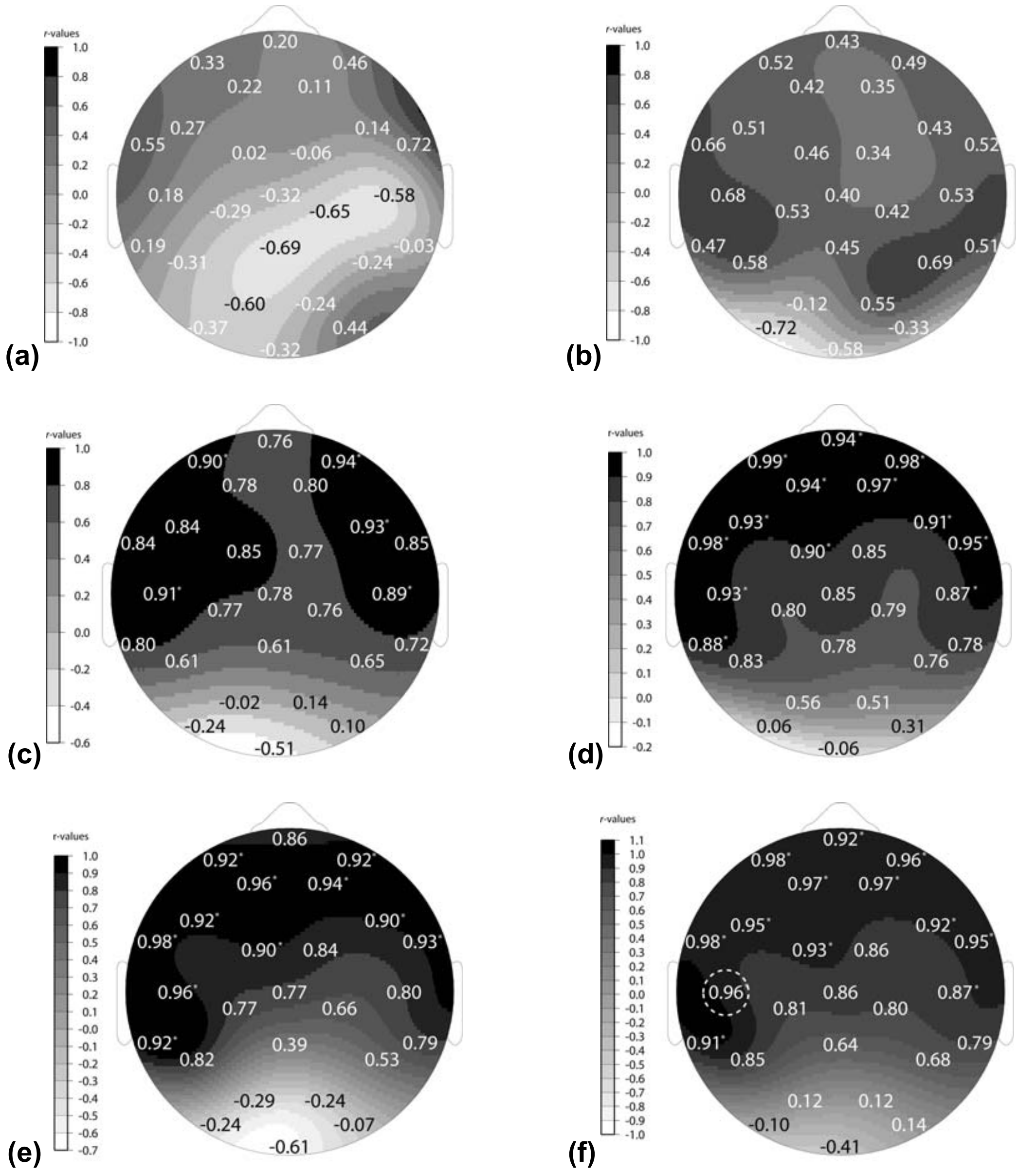


Figure 4.19(a-f). Correlations of "traditional" constraint measure with UA-UN noun mean amplitude over several time windows: (a) N1 (50-150 ms), (b) P2 (150-250 ms), (c) N400 (200-500 ms), (d) Early LP (500-800 ms), (e) Late LP (800-1200 ms), and (f) Extended LP (500-1200 ms). Significant correlations (p ≤ .05) are indicated with asterisks (*).

(g)

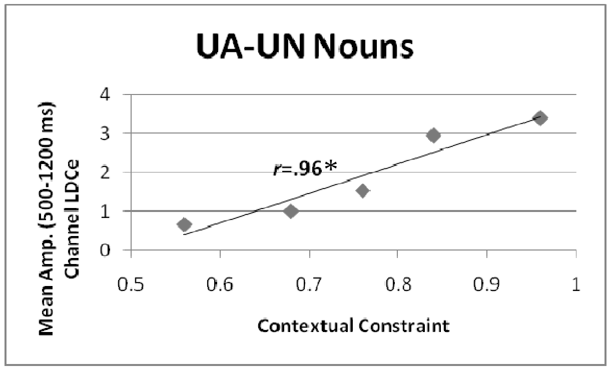


Figure 4.19(g). Scatterplot of UA-UN noun correlations over Extended LP time window (500-1200 ms) at a single electrode site (LDCe) highlighted in Figure 4.19(f).

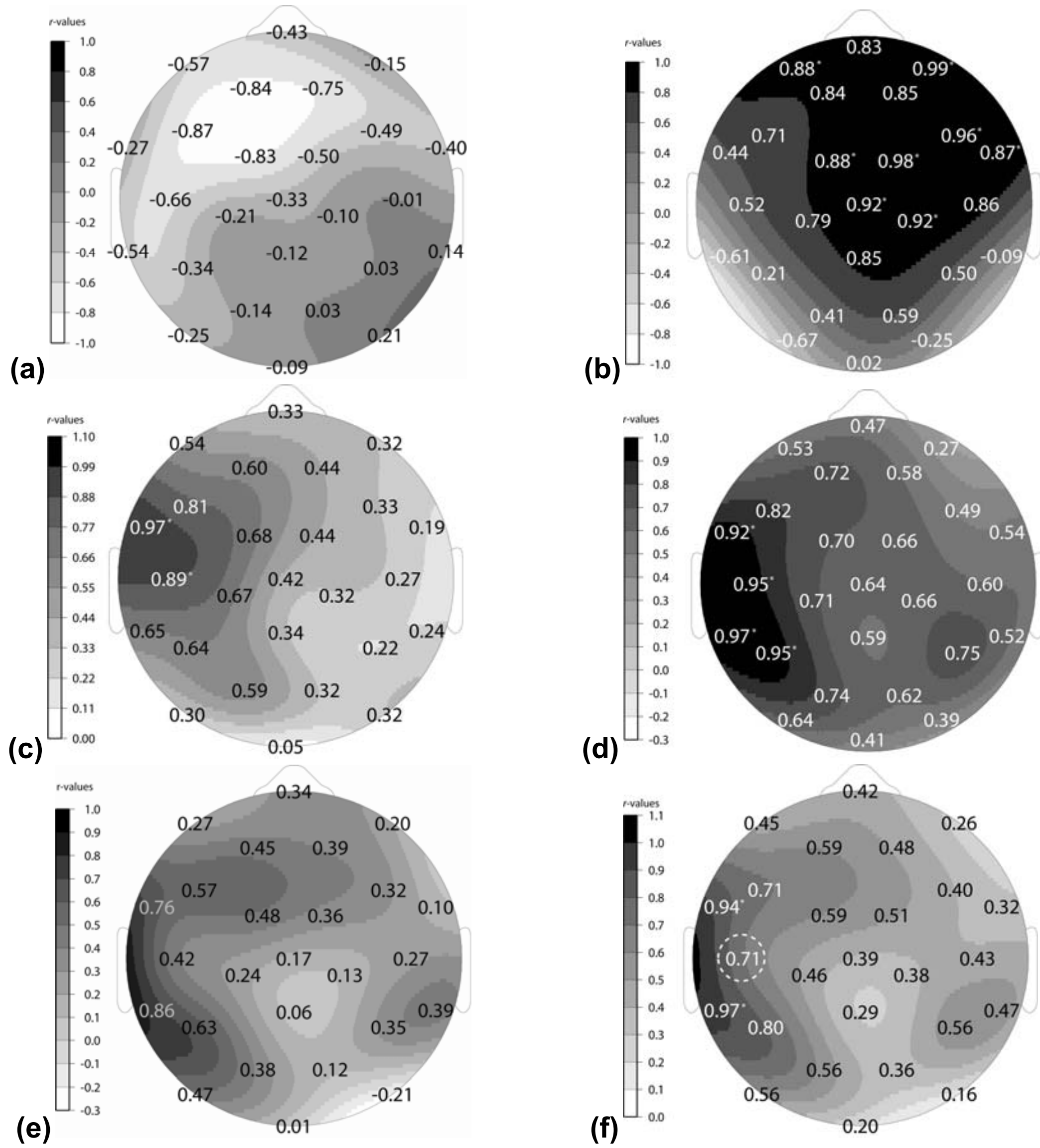


Figure 4.20(a-f). Correlations of "traditional" constraint measure with EA-UN noun mean amplitude over several time windows: (a) N1 (50-150 ms), (b) P2 (150-250 ms), (c) N400 (200-500 ms), (d) Early LP (500-800 ms), (e) Late LP (800-1200 ms), and (f) Extended LP (500-1200 ms). Significant correlations (p ≤ .05) are indicated with asterisks (*).

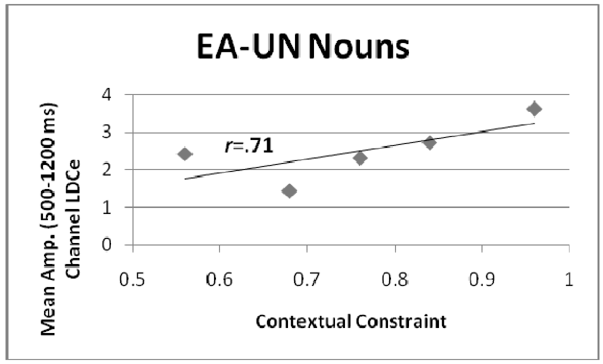
(g) 

Figure 4.20(g). Scatterplot of EA-UN noun correlations over Extended LP time window (500-1200 ms) at a single electrode site (LDCe) highlighted in Figure 4.20(f).

4.11.1.1.1. N1 time window

Over the 11 frontal electrode locations where the N1 measurements were evaluated, there did not appear to be a consistent pattern of correlation values for the UA-UN nouns (**Figure 4.19a**), with only a single electrode site (RLFr) exhibiting an r-value that approached significance (*r*=0.72). For EA-UN nouns (**Figure 4.20a**), contextual constraint and N1 amplitude showed strong inverse correlations over left medial electrode sites, with N1 negativity increasing as the degree of constraint violation increased. Correlation values were maximal at electrodes LDFr (*r*=-0.87) and LMPf (*r*=-0.84).

4.11.1.1.2. P2 time window

Using the 11 frontal electrodes where the frontal P2 effect was measured, the P2 mean amplitude of UA-UN nouns did not appear to correlate significantly with contextual constraint violation (**Figure 4.19b**). The P2 of the EA-UN nouns (**Figure 4.20b**), however, was strongly correlated with contextual constraint violation: as the degree of constraint violation increases, so does P2 amplitude, with maximal correlation values at RLPf (*r*=0.99) and RMFr (*r*=0.98).

4.11.1.1.3. N400 time window

As part of ascertaining a double dissociation between the effects of cloze probability and constraint violation on N400 and LP amplitude, we first sought to verify – as has been observed in previous experiments– that N400 amplitude does not vary as a function of constraint violation. For the UA-UN nouns, **Figure 4.19c** reveals a pattern of positive correlations (except at three posterior sites), highest at anterior sites (maximal r= 0.94 at RLPf). These positive *r*-values indicate that as constraint violation increases, mean ERP amplitude in the N400 time window becomes more positive. Similarly, correlations for the EA-UN nouns (**Figure 4.20c**) also exhibit a pattern of positive correlations, with maximal *r*-values clustering at left central sites, instead of over scalp locations more typical for N400 effects (i.e., centro-parietally, rightish). These analyses reveal that indeed, N400 amplitude does *not* increase as a function of contextual constraint violation; rather, the high positive correlation pattern and its left/frontal scalp distribution is indicative of the presence of an LP beginning already in this earlier time window.

4.11.1.1.4. Extended LP time window (500-1200 ms)

For both UA-UN and EA-UN nouns, the pattern of (mostly) positive correlation values at individual electrode channels indicates that the higher the sentence constraint, the larger the late positivity in the ERP when participants received an unexpected noun following the expected article for those same contexts. For the UA-UN nouns in **(Figure 4.19f)**, the distributional pattern of *r*-values indicates that the positivity is most highly correlated with violations of sentence constraint over anterior sites (maximal *r*=.98 at left frontal/prefrontal sites, with the minimum *r*-value at any of the 13 frontal sites being .86), with slightly higher values over the left than right hemisphere. For the EA-UN nouns **(Figure 4.20f)**, which were also highly correlated with constraint violation (maximal *r*=.97 at left temporal site), the pattern of high *r*-values is more focal, with maximal values clustering over left, central recording sites.

4.11.1.1.5. Early (500-800 ms) and Late (800-1200 ms) LP time windows

Visual inspection of the correlations for both Unexpected noun types using the Early (**Figures 4.19d** and **4.20d**) and Late (**Figures 4.19e** and **4.20e**) LP time windows seemed to indicate very similar correlation strengths and distributional patterns as the Extended (**Figures 4.19f** and **4.20f**) LP time window. However, in order to compare the spatial distributions of the LP correlations in the early (500-800 ms) and late (800-1200 ms) LP time windows for the two types of Unexpected nouns, repeated measures ANOVAs with subjects as a random factor and the factors time window (Early LP, Late LP), hemisphere (left, right) and anteriority (anterior, posterior) were conducted for each noun type using 5 electrode sites per scalp quadrant (20 locations total), as shown in **Figure 4.21**, below.

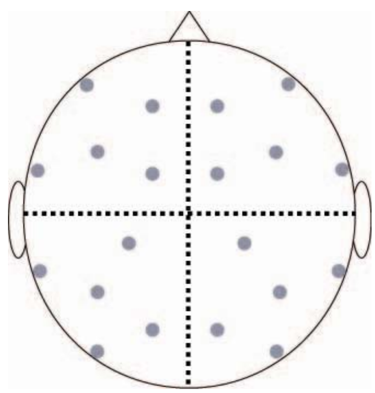


Figure 4.21. Locations of the 20 electrode sites, divided into four quadrants, used to compare distributional patterns between different correlation analyses.

Fisher's *r*-to-Z transformations were calculated to express the correlation values in units normalized by their standard deviation (Z-scores). For the UA-UN nouns, there was no main effect of time window [$F(1,1) = 2.47$, $p = 0.13$, ns], but there was an overall effect of anteriority, with higher positive correlations at anterior than posterior sites [$F(1,1) = 47.58$, $p < 0.0001$]. For the EA-UN Nouns, the correlations in both LP time windows for the EA-UN nouns were biased toward the left hemisphere [$F(1,1) = 18.37$, $p = 0.0002$]. There was also a main effect of time window, with significantly larger correlations in the Early LP time window [$F(2,2) = 20.29$, $p < 0.001$] than in the Late LP time window.

4.11.2. Alternative contextual constraint analysis

As mentioned, contextual constraint is most typically operationalized as the cloze probability of the most expected continuation for a particular sentence. However, with respect to the stimuli used in the current study, there is another way to think about how constraint could be quantified, which is to consider how strongly comprehenders hold onto their original noun expectations in the face of an Unexpected indefinite article (**Figure 4.22**).

"Alternative" contextual constraint measure for UA-UN nouns. Determined by norming results with Unexpected articles:

e.g., *Dale was very sorry and knew he owed Mary a...*

where *apology* is the Expected continuation for the context normed with the expected indefinite article *an* (cloze of *apology* = 97%),

but *check* (in the UA-UN condition) is the actually presented noun in the ERP experiment.

Norming results:

Adjective strategy: (e.g., *sincere, major, heartfelt, huge, personal*) + *apology* (cloze of *apology* preceded by any adjective = **84%**) ← **"Alternative" constraint determined by this value**

Alternate noun strategy: e.g., *kiss, sorry, dollar* (cloze of all noun responses =13%)

Figure 4.22. Alternative constraint measure (for UA-UN nouns only), sample stimulus.

To investigate effects of this alternative formulation of constraint violation, we again turned to our stimulus norming results, this time examining the responses to sentences normed with the *Unexpected* articles. For instance, for the sentence context in

Figure 4.22, 84% of norming participants supplied an adjective (not any one particular adjective) followed by the Expected noun. Under our alternative quantification of constraint, this specific context would be fairly highly constraining (84%). However, the different sentence stimuli varied in the extent to which participants used this strategy to continue the sentences versus supplying an alternative noun. For instance, an example of a sentence context for which providing an alternative noun was the predominant strategy is the following:

(7) *The student had asked her question five times already. She was still waiting for a ...*

where *answer* is the Expected continuation when the context is normed with the expected indefinite article *an* (cloze of *answer* = 97%),

but *letter* (in the UA-UN condition) is the actually presented noun in the ERP experiment.

Norming results for sentence context (7):

- Adjective strategy (any adjective + the expected noun) = 34% cloze, e.g., (*detailed, reliable, thorough, decent, proper*) + *answer*
- Alternate noun strategy (a different single noun response) = 53% cloze, e.g., *response*

In sentence context (7), constraint in terms of maintaining the original noun expectation by way of an adjective is 34%. To determine a constraint value for each context, then, we examined the values for each sentence stem normed with the Unexpected article and calculated the cloze probability of the Expected noun preceded by *any* adjective, e.g., 84% for the sentence in **Figure 4.22** and 34% for context (7) above. Again, this method deviates from the way in which constraint is typically determined – i.e., calculating constraint as the cloze probability of the most commonly provided (usually single word) continuation for the context normed with the Expected article. Our rationale for using this alternative method was to ascertain how much or how little the presentation of the indefinite article may have altered the expectation for the original (pre-activated) noun. Based on these values, we then sorted the sentence stimuli into ten 10% bins spanning the interval 0-100% and analyzed the ERPs for the UA-UN nouns (e.g., *check* or *letter*) for the respective bins. In each such bin, the mean ERP amplitude of the UA-UN nouns as well as the mean cloze probability for an adjective plus the Expected noun (the constraint), were computed. It is important to note the cloze probabilities of the actually presented UA-UN noun bins were all similarly low (see **Table 4.4**). Correlations were then computed at each electrode site using the mean alternative constraint value of each bin and the mean UA-UN noun ERP amplitude for each bin. Note that this alternative method for calculating contextual constraint cannot be performed for the EA-UN nouns, as there are no prediction-consistent indefinite articles to force any sort of change in expectations.

Table 4.4. Alternative constraint analysis breakdown of constraint conditions.

Constraint Bin Using Adjective Strategy	Number of Items per Bin	Average Bin Constraint	Average cloze of presented UA-UN nouns
90-100%	12	.92	.01
80-89%	17	.84	.00
70-79%	18	.73	.02
60-69%	20	.64	.03
50-59%	30	.55	.03
40-49%	20	.46	.02
30-39%	18	.34	.04
20-29%	12	.25	.03
10-19%	10	.17	.07
0-9%	3	.07	.00

4.11.2.1. Correlations of alternative constraint with ERP components' mean amplitudes

Similar to our correlation analyses using traditional contextual constraint values, we performed correlations of alternative contextual constraint with UA-UN mean amplitude over six different time windows: N1 (50-150 ms), P2 (150-250 ms), N4 (200-500 ms), early LP (500-800 ms), late LP (800-1200 ms), and extended LP (500-1200 ms). See **Figure 4.23a-g**.

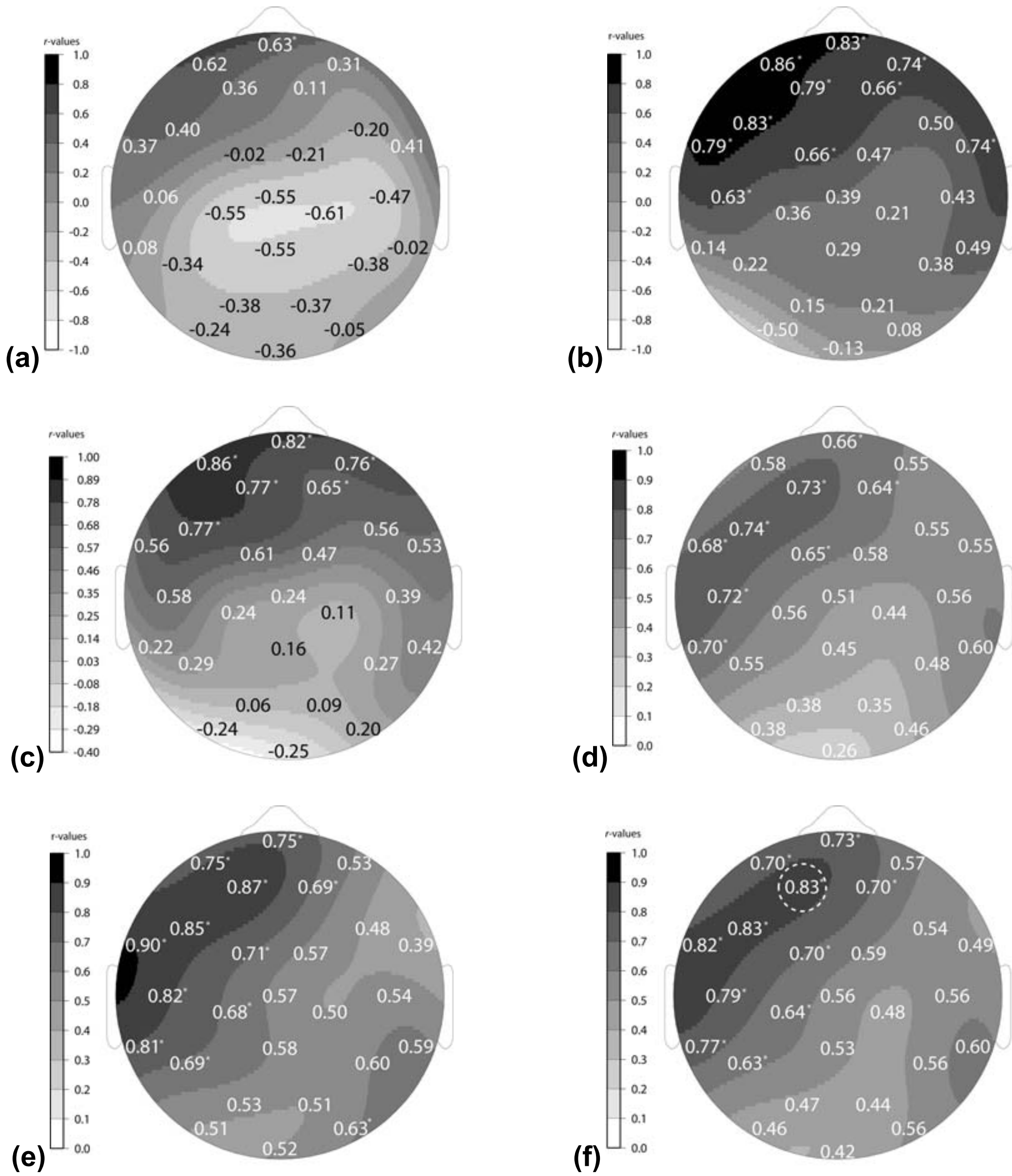
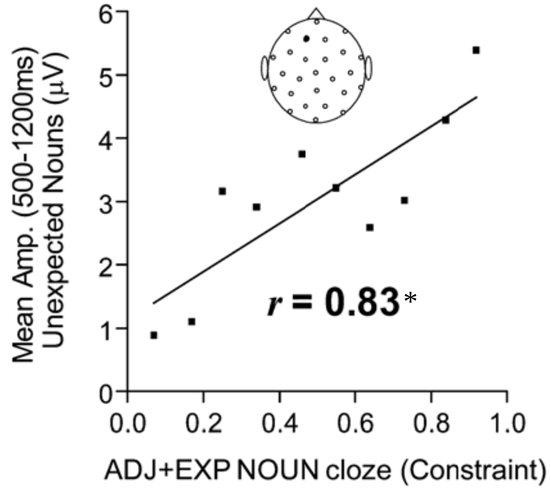


Figure 4.23(a-f). Correlations of "alternative" constraint measure with UA-UN noun mean amplitude over several time windows: (a) N1 (50-150 ms), (b) P2 (150-250 ms), (c) N400 (200-500 ms), (d) Early LP (500-800 ms), (e) Late LP (800-1200 ms), and (f) Extended LP (500-1200 ms). Significant correlations (p ≤ .05) are indicated with asterisks (*).



(g)

Figure 4.23(g). Scatterplot of UA-UN noun correlations over Extended LP time window (500-1200 ms) at a single electrode site (LMPf) highlighted in Figure 4.23(f).

4.11.2.1.6. N1 time window

At none of the 11 frontal electrode sites where the N1 was assessed were there significant correlations of constraint violation with ERP mean amplitude (**Figure 4.23a**).

4.11.2.1.7. P2 time window

Using the same 11 frontal electrodes as the N1 analysis, P2 amplitude and constraint violation were highly positively correlated, particularly at left frontal scalp locations (maximum r=0.86 at LLPf): P2 amplitude increased with constraint violation (**Figure 4.23b**).

4.11.2.1.8. N400 time window

Again, with the goal of determining a double dissociation between the effects of cloze probability and constraint violation on N400 and LP amplitude, we tested whether N400 amplitude varied as a function of constraint violation. **Figure 4.23c** reveals a pattern of positive correlations (except at two posterior sites), highest at anterior sites (maximal r= 0.86 at LLPf), indicating that as constraint violation increases, mean ERP amplitude in the N400 time window becomes more positive (the opposite direction of what might be expected if N400 negativity increased with constraint violation). These analyses confirm that similar to the traditional constraint analyses, N400 amplitude is not modulated as a function of contextual constraint violation.

4.11.2.1.9. Extended LP time windows (500-1200 ms)

Figure 4.24a shows a subset (3 of 10) of the constraint bins at a single frontal electrode channel, MiPf, and a bar graph (**Figure 4.24b**) of the mean amplitudes of these bins over prefrontal, central and occipital midline electrodes. These patterns indicate that the more likely the norming subjects were to hold onto their expectations for Expected nouns by means of adjectives when they received the Unexpected article, the greater the LP between 500-1200 ms when the ERP participants received an Unexpected noun following the Unexpected article for those same contexts.

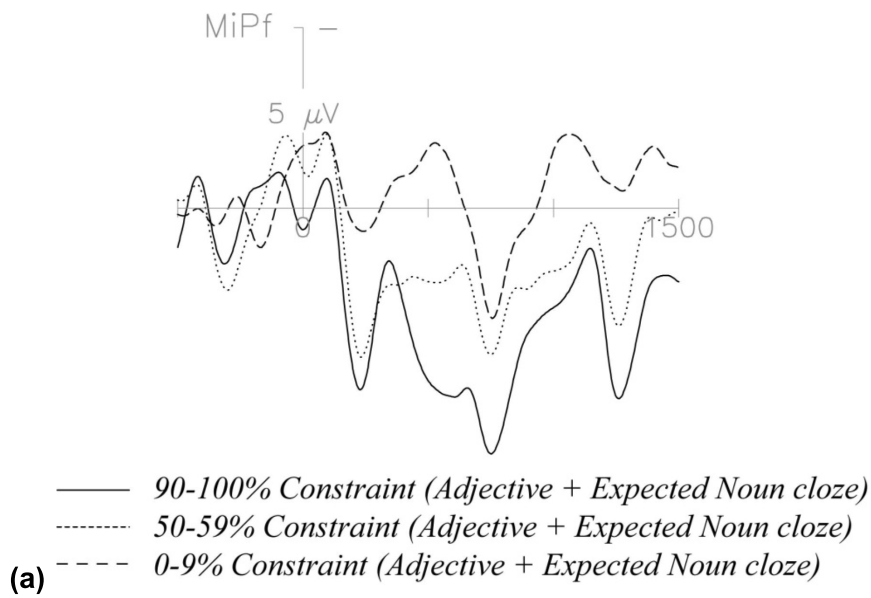
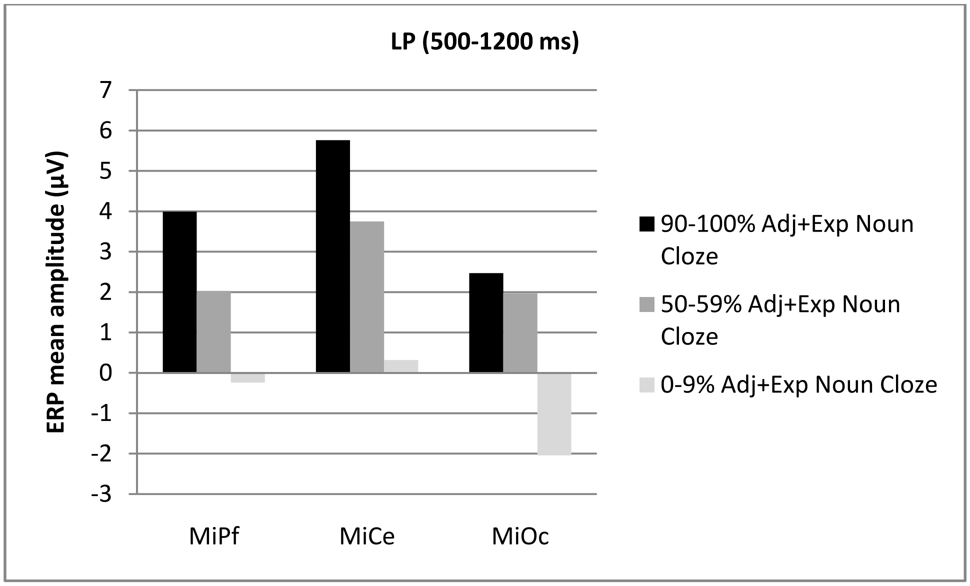


Figure 4.24(a). Low pass-filtered UA-UN nouns sorted by ADJECTIVE+EXPECTED NOUN cloze for respective sentence contexts when unexpected article is provided at a single electrode site (MiPf).



(b)

Figure 4.24(b). A bar graph of the mean amplitudes of the bins in Figure 4.24(a) over prefrontal, central and occipital midline electrodes.

The LP correlation pattern that emerged for the UA-UN nouns when the alternative measure of sentence constraint was utilized appeared similar to the pattern observed for the correlations using traditional constraint analysis: maximal correlations (*r*=.83) at frontal sites, albeit with an apparent left hemisphere bias for the alternative constraint correlations (see **Figure 4.23f,g**). As with the traditional constraint correlations, the LP to the Unexpected noun systematically increased as constraint violation increased.

4.11.2.1.10. Early (500-800 ms) and Late (800-1200 ms) LP time window

Visual inspection of the correlations for UA-UN nouns using the Early (**Figure 4.23d**) and Late (**Figure 4.23e**) LP time windows seemed to suggest similar scalp distribution patterns, but with slightly higher correlations in the Early LP time window. To directly compare the spatial distribution of the LP correlations in the two time windows, we performed a repeated measures ANOVA identical to that conducted for the analyses of Traditional constraint. Our results revealed main effects of hemisphere and anteriority (left and frontal LP biases), mediated by an interaction of the two factors [$F(1,1) = 6.60$, $p = 0.02$], indicating a larger left bias over frontal than posterior sites. There was a main effect of time window, [$F(1,1) = 9.66$, $p < 0.01$] indicating that correlations were significantly lower in the Early LP time window than in the Late LP time window; however, this effect interacted with hemisphere [$F(2,2) = 5.31$, $p = 0.03$], such that there was a larger left bias in the Late LP time window than in the Early LP time window.

4.11.3. LP: Comparison of traditional and alternative contextual constraint analyses

To compare the LP correlations observed across all three contextual constraint analyses (Traditional constraint for both unexpected noun types as well as Alternative constraint for UA-UN nouns), an omnibus ANOVA on the Fisher *Z*-transformed correlations was performed using 3 levels noun type/constraint analysis (UA-UN Traditional, EA-UN Traditional, UA-UN Alternative) X 2 levels hemisphere (left, right) X 2 levels anteriority (anterior, posterior) X 2 levels time window (Early LP, Late LP). Results indicated that overall, there was a left, frontal, and early LP time window bias across all three correlation analyses, with significantly larger effects for the Traditional Constraint UA-UN analysis than for the other two analysis types (see **Table 4.5**). These main effects were mediated by interactions of noun type/analysis X anteriority, revealing that the frontal bias emanated from the two UA-UN analyses [$F(2,2) = 29.12$, $p<0.0001$]. Finally, there was an interaction of Noun Type/Analysis with time window [$F(2,2) = 8.08$, $p<0.0006$], showing an Early LP time window bias (500-800 ms) for the two Traditional Constraint analyses, but larger effects over the Late LP time window (800-1200 ms) for the Alternative constraint analysis.

Table 4.5. Comparison of LP effects across different constraint analyses.

Analysis Factors	UA-UN Nouns, Traditional Constraint Analysis	EA-UN Nouns, Traditional Constraint Analysis	UA-UN Nouns, Alternative Constraint Analysis
Noun Type/Constraint Analysis	Strongest overall correlations		
Hemisphere	Left bias	Left bias	Left bias
Anteriority	Frontal bias	No anteriority bias	Frontal bias
Time Window	Early LP (500-800 ms) bias	Early LP (500-800 ms) bias	Late LP (800-1200 ms) bias

4.12. Sentential contextual constraint analyses: Discussion

Several of the tested time windows appeared to show sensitivity to degree of constraint violation. Our analyses using variable measures of contextual constraint are summarized below by ERP effect of interest (Table 4.6).

Table 4.6. Constraint correlation analyses grouped by type of constraint measure and ERP effect/noun time window. Significant correlations (p ≤ .05) are indicated with asterisks (*).

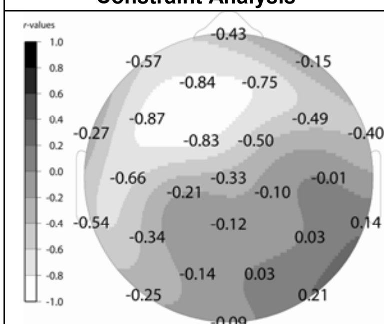
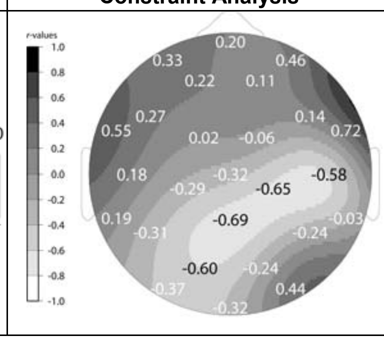
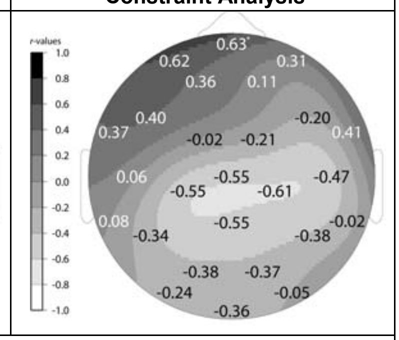
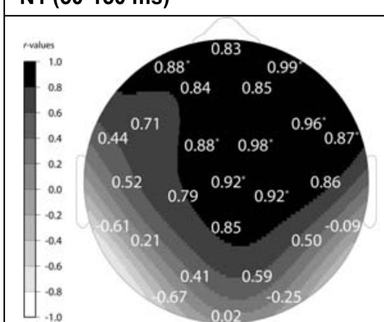
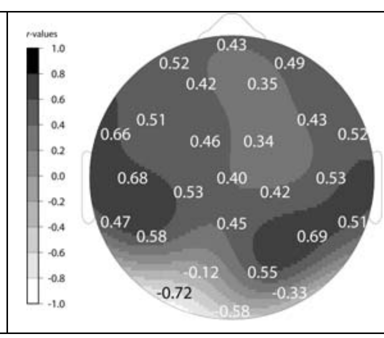
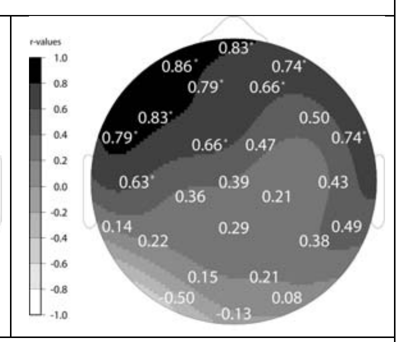
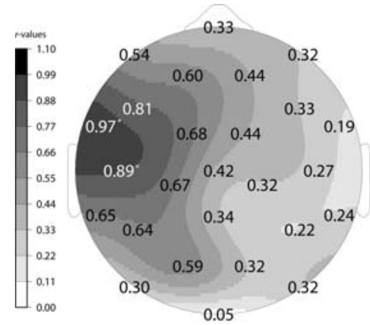
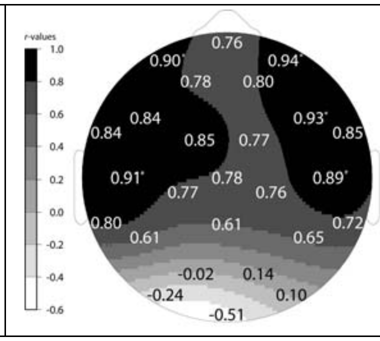
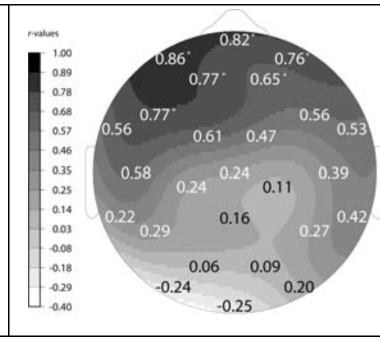
EA-UN "Traditional" Constraint Analysis	UA-UN "Traditional" Constraint Analysis	UA-UN "Alternative" Constraint Analysis
		
N1 (50-150 ms)		
		
P2 (150-250 ms)		

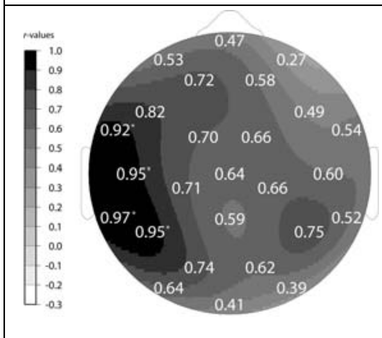
Table 4.6 continued

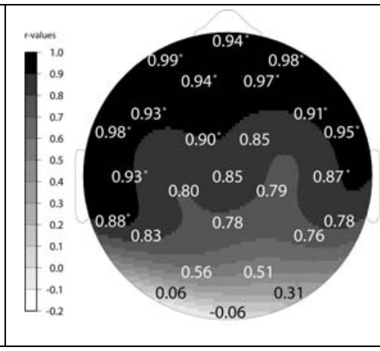


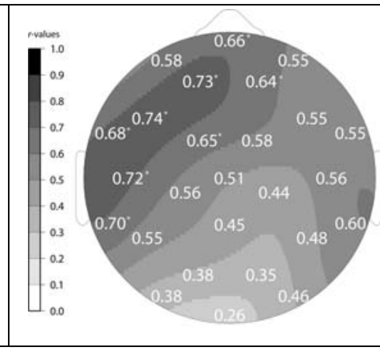




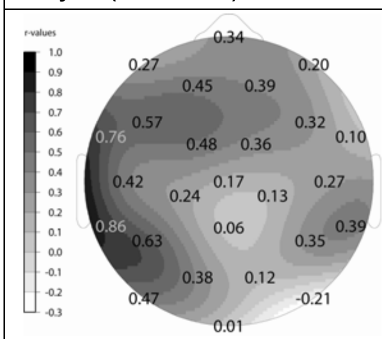
N400 (200-500 ms)

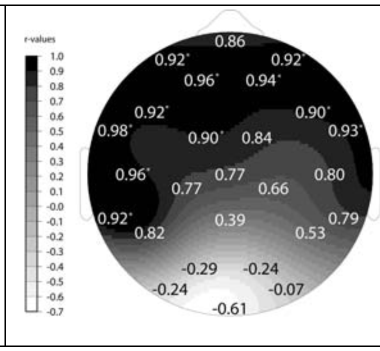


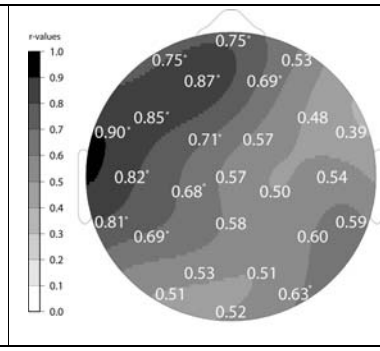




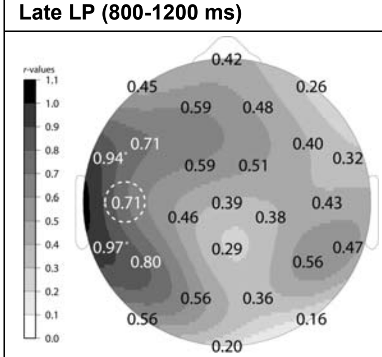
Early LP (500-800 ms)

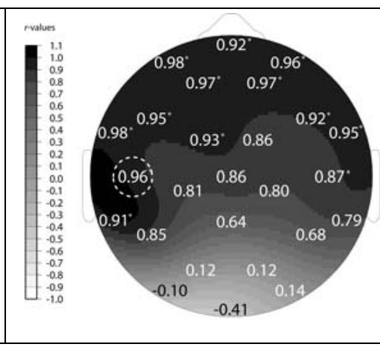


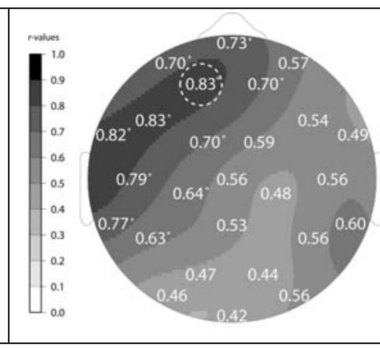




Late LP (800-1200 ms)







Extended LP (500-1200 ms)

4.12.1. Frontal N1

Only the traditional constraint analysis of the EA-UN nouns yielded a pattern of significant negative correlations: as the degree of constraint violation increased, ERP amplitude became more negative (N1 amplitude increased).

4.12.2. Frontal P2

Analyses using traditional measures of constraint indicated that the P2 amplitude of EA-UN, but not UA-UN, nouns increased as a function of constraint violation. Using an alternative constraint measure, however, UA-UN nouns exhibited a similar correlation pattern.

4.12.3. N400

While there were effects of constraint violation within the N400 time window for both UA-UN and EA-UN nouns regardless of the type of constraint analysis, these effects a) were *opposite* of what would be predicted if N400 amplitude and constraint violation were related (i.e., ERP amplitude became more *positive* with increasing constraint violation), and b) were maximal at left and frontal electrode sites (scalp locations not canonically associated with the N400 component). This suggests, then, that what we refer to as the "Late Positivity" may actually begin earlier in the epoch than we had originally thought. We propose that there may be some overlap in the timing of the posterior N400 and the more anterior LP which makes it difficult to isolate the precise onset point of the LP. Our finding that the N400 is insensitive to constraint violation confirms results of previous studies (e.g., Kutas & Hillyard, 1984) that evaluate this relationship more directly. These results also speak to one half of the N400-LP double dissociation puzzle: that is, the N400, which correlates highly with the cloze probability of a particular item, is *not* correlated with sentential constraint violation.

4.12.4. Late positivity (LP)

Using two different measures of sentential constraint, we showed that the LP first observed to Unexpected relative to Expected nouns in our original cloze analysis is sensitive to the degree that Unexpected nouns violate contextual constraint. While cloze probability and constraint are parallel measures in the upper range, the two dissociate in the lower range; that is, you can have an item with only 2% cloze probability (unexpected) as a continuation in either (a) a highly constraining context or (b) a not at all constraining context. Though the amplitude of the N400 component to unexpected continuations in (a) and (b) is undifferentiated, we demonstrate here that another component, the LP, *is* sensitive to the difference between (a) and (b). The LP, however, does not seem to serve as an index of cloze probability per se. (Though an LP difference was observed in the analysis of Expected versus Unexpected nouns when the nouns were collapsed across constraint, there was no systematic relationship when cloze and LP mean amplitude were correlated.) We will argue in the General Discussion section that the double dissociation demonstrated herein provides strong support for a predictive language processing model by identifying a graded “cost” to violations of constraint.

In our analysis using a traditional measure of contextual constraint, LP amplitude increased linearly as function of how strongly the most expected continuation was anticipated. Over the Extended LP time window (500-1200 ms), the maximal correlations for both UA-UN and EA-UN nouns were similarly high (r=.98 and .97 maxima, respectively) with a left bias for both noun types. For unknown reasons the pattern of LP correlations were generally higher, more frontal, and less focal for the UA-UN nouns. One possibility may be that for the UA-UN nouns, the Unexpected noun is signaled in advance by the Unexpected article, and that may serve to enhance/engage more processing resources for whatever function the LP may be indexing, though this is merely speculation on our part.

The use of the alternative measure of contextual constraint to analyze the UA-UN nouns resulted in a correlation pattern similar in strength and distribution (with correlations between constraint violation and LP amplitude highest in the left anterior scalp quadrant) to that observed using the traditional constraint values. This is not so surprising, after all, in that both measures reflect levels of expectation for (for the most part) highly expected nouns: the difference being that the alternative constraint measure takes into account that sometimes the Unexpected article forces a change in what the most expected continuation will be. Given the correlation results using the alternative constraint measure, it is clear that the articles are treated by the brain as relevant information, and are not simply ignored during on-line sentence comprehension. However, even in the face of an Unexpected article, the brain's parser may still choose to keep a highly anticipated continuation activated, under the possibility that the item may still be received as input somewhere "down the line".

One question spurred by these results is the degree to which the LP ERP effect detected to constraint violations is similar to other prolonged LPs, including those observed in traditional semantic N400 studies (e.g., with manipulations of expected vs. unexpected, plausible vs. implausible sentence continuations), to the variety of manipulations in recent "semantic P600" studies (e.g., with the component linked to such functions as thematic role assignment, subcategorization bias, and animacy violation), and to more traditional P600s to syntactic manipulations (e.g., agreement violations and structural ambiguities). In addition to the ever-present inverse problem, the variability in scalp distributions and precise latencies between studies make this a tricky issue, especially since P600 effects have historically encompassed a wider range of timing and topography than, say, N400s (Groppe, 2008). Our finding of an enhanced LP to Unexpected nouns in constraining contexts revealed through correlations, was most evident at left anterior scalp sites over a 500-1200 ms time window; thus, our effect may be more frontal and more prolonged than "traditional" P600 findings, but still within the range of what has been labeled as such. Indeed, P600s elicited by studies using syntactic violation manipulations have often been found to have a more posterior scalp distribution than those elicited by studies manipulating syntactic ambiguity (Osterhout & Hagoort, 1999). Kaan & Schwaab (2003) suggest distinctly posterior and frontal P600s, which index syntactic processing difficulty and ambiguity resolution/increased discourse level complexity, respectively. Ultimately, the main goal of the current study is not to attach a label to the ERP effect we have observed, or for that matter derive a comprehensive theory that functionally unites (or dissociates) all the above-mentioned LP findings. For our purposes, we would like to understand the sensitivities of the effect to the extent that we are able to interpret how on-line language proceeds; in particular, we would like to understand whether anticipatory language processing is part-and-parcel of how the comprehension system works, and if it is, whether there is a "cost" to a predictive system that cannot possibly consistently and accurately anticipate/pre-activate upcoming material. If our proposal of an LP effect with a sensitivity to contextual constraint is correct, then this could be exploited for future on-line language comprehension studies, though in the end it could also serve as another piece in the twin puzzles of determining the functional nature of the P600 and theorizing about the domain specificity (syntax versus semantics, or some kind of continuum) of language processing. We will explore the theoretical implications of our LP's amplitude modulation as a function of constraint violation further in the General Discussion section.

4.13. General discussion

Our results did not support our original hypothesis, that there may be a cost to prediction violation that is specifically syntactic in nature. The LP following the N400 for

Unexpected nouns was elicited whether an Unexpected noun followed a high cloze (prediction-consistent) or low cloze (prediction-inconsistent) indefinite article. Thus, there was no indication that the prediction-inconsistent article forced comprehenders (on-line) to expect an adjective and consequently experience processing difficulty with the presentation of a different part of speech (a subsequent Unexpected noun). However, upon reanalysis, our results did support the idea that there is a processing cost to constraint (prediction) violation and that this cost is reflected in a late positive-going ERP deflection. Additionally, we demonstrated that the N400 and the LP offer separate indices of word expectancy (cloze, reflecting the probability of an actual item, regardless of prior contextual constraint) and contextual constraint violation (the difference between the degree to which a context constrains for a particular item and the probability of the actually presented item) during sentence processing. Our analyses of the contextually Unexpected noun ERPs indicated that the amplitude of the LP was graded based upon the strength of the expectation for the highest cloze noun for that context (aka sentence constraint). Using two measures of constraint which had in common that they both utilized the cloze probability of the Expected but not presented noun, the mean amplitude of the LP to the Unexpected nouns was found to be highly correlated over certain scalp areas with the degree to which sentential constraint had been violated. The more constraining a context was, the larger the LP when an Unexpected noun continued that context. Using a more traditional quantification of contextual constraint (i.e., the cloze probability value of the most frequent sentence continuation for the context), for the UA-UN nouns, there was a left hemisphere, frontal, Early (500-800 ms) bias for the broadly distributed LP-effect. The EA-UN nouns were also highly correlated with constraint violation, however the scalp distribution of highly significant *r*-values was more focal, but with maximal values still clustering over left hemisphere recording sites during the Early LP time window. A similar LP pattern emerged for the UA-UN nouns when an alternative measure of sentence constraint (i.e., constraint determined by the how strongly the expectation for the Expected noun was preserved in the face of receiving an Unexpected indefinite article) was utilized: maximal correlations were at frontal sites, again with a bias towards the left hemisphere. For the alternative constraint analysis, though, stronger correlations were noted in the Late (800-1200 ms) rather than Early LP time window.

What to make of the distributional and timing differences in the late positivity as a function of the different constraint analyses? In effect, each of the three analyses measured a slightly different aspect of constraint violation. For instance, because the Traditional constraint measure is based upon the cloze of the most probable continuation when contexts are normed with the *Expected* article provided, the implications for analyzing the UA-UN versus EA-UN are slightly different; in particular, for the UA-UN nouns this means the measure essentially ignores the influence of the Unexpected article on constraint. When contrasting the two noun types, this also means that the UA-UN nouns have been "pre-signaled" by their Unexpected articles such that whatever functional processes the LP is indexing may be initiated earlier/later, be more extended/compressed in time, overlap to a greater/lesser degree with other components, or perhaps differ in strength. The alternative constraint analysis of the UA-UN nouns, on the other hand, explicitly takes into account the influence of the Unexpected article on constraint and measures the degree to which comprehenders' offline "recovery" strategies map onto the brain's online response to these stimuli. This alternative analysis also resulted in a wider range of constraint measures, from 0-100% rather than the 50-100% range in the traditional constraint analysis, which is also very likely to have contributed to the difference in correlation patterns. Had this experiment been designed from the outset to test constraint violation directly, we surely would have included a greater number of lower constraint contexts, so that a full 0-100%

range of constraint could be utilized, which may in turn have resulted in more parallel results between analyses.

However, our observation of an enhanced (consistently) left and fronto-centrally-biased LP that systematically varies with the degree of constraint violation adds another piece to the growing, but somewhat complex, pattern of recent literature findings of semantic LPs. While our study does not directly address what the precise functional correlate of the component might be, it may not be incompatible with similar late positive-going EPs that have been elicited by non-syntactic manipulations: we will return soon to a comparison of our results with some of these. The first point we will take up, though, is that our constraint violation LP is relevant for the purposes of showing that the brain is using context to predictively generate expectancies for upcoming linguistic items and that there is a processing cost when strongly pre-activated items are not received.

4.13.1. Support for predictive language processing and a cost for misprediction

In Experiment 1 we argued on the basis of graded N400s at English indefinite articles preceding more and less expected target nouns, that the parser predictively activates possible continuations based on representations of the sentence context that have accrued up until that point, and that integration is less facilitated (as indexed by larger N400 amplitude at the article) when it receives input that disconfirms/is not compatible with the most likely noun continuation. In the same study, a post-N400 LP was also observed to contextually less expected nouns that followed less expected indefinite articles. The current experiment was conducted to test whether this LP, in line with extensive findings of P600s in response to syntactic violations or ambiguities, was similarly reflecting the violation of some kind of syntactic expectation for a particular part of speech (an adjective) as a consequence of receiving an unexpected indefinite article. To test this hypothesis, we used sentences (with a limited range of constraint) with unexpected target noun continuations that were preceded by either less or more contextually expected indefinite articles. For the syntactic expectancy hypothesis to have held, LPs should have been observed only at the Unexpected target nouns preceded by Unexpected pre-nominal articles, but not at the target nouns preceded by an expectancy-compatible article. However, both types of Unexpected nouns elicited LPs – a finding that was compatible with more general (i.e., non-syntactocentric) violation of contextual constraint. Furthermore, the amplitude of the LP to Unexpected nouns (with low cloze probability held relatively constant) was found to increase as a function of sentence constraint for an expected, but not presented, item.

On the face of things, the possibility of a "cost" or "cost gradient" is compatible with the findings from Experiment 1, which indicated that ERP effects of prediction were probabilistic: even for less constraining sentences, the N400 showed some sensitivity to the low level of pre-activation for particular target items. This graded prediction effect, in which N400 amplitude varied linearly in relation to the cloze probability of the eliciting item, suggested that there is some pre-activation of potential continuations even when the preceding context did not extensively constrain the number of possible continuations. This would imply that there is always *some* degree of expectation to be dealt with when the parser encounters an unexpected item, whether that expectation is a strong one, or a relatively weak one. In turn, one might expect this graded relationship of the N400 as an index of prediction to have implications for how the parser recovers from situations in which expectations are violated by the presentation of low cloze probability items. It is thus perhaps not so surprising that in the current study we found a graded LP effect to unexpected nouns that varied as a function of the strength of expectancy (constraint) for the most predictable items.

We would like to argue that our finding of strong correlations of the LP with degree of constraint violation is in line with a predictive model of sentence processing. While our

Experiment 1 results found evidence for prediction of lexical forms prior to the actual target words, the current results support the view that there is a consequence to highly pre-activating upcoming items but not receiving them, (in contrast to the context *not* having strongly pre-activated possible continuations.) Additional processing, as indexed by the LP, seems to be present for the first case, but not (or at least not to the same degree) for the second. Under a strictly incremental model of sentence processing, one could imagine that for both strongly and weakly constraining contexts there might be processing difficulty (as indexed by the N400) when it came to integrating a low cloze continuation into the contextual representation that had been formed up until that point. But under such a model, what explanation could there be for the additional processing indexed by the semantic LP effect when there has been a violation of constraint? Until recently, most views of the P600 have linked the component to syntactic reanalysis, revision or repair. Under such views, it is possible to consider that an incremental parser may be assigning (based on such phenomena as case marking, animacy, thematic roles, or subcategorization frames, etc.) certain structure to the accruing context, such that an unexpected item may cause, for instance, thematic roles to have to be reassigned or a different parse of the syntactic structure to be pursued. This is not, however, the case with the stimuli in the current study. In our study, the Unexpected nouns varied from the Expected ones only in their contextual cloze probabilities (e.g., semantic, categorical, lexical fit, or with respect to event/word knowledge) but did not alter the representation of the syntactic structure, thematic roles, etc. of the sentence constructed up to that point, ruling out a syntactic explanation of our LP finding. However, the fact that the graded LP effect at Unexpected nouns was present to a greater degree when sentence context was highly constraining, strongly supports the idea that contextual information was used to form predictions about upcoming input, and when the expected input was not received, some additional processing was required.

Our findings, in which we have attempted to link the systematic variation in post-N400 LP amplitude to violation of sentential constraint are not without precedent. In particular, the positivity in our study (between 500-1200 ms) most closely maps onto a frontal positivity (500-900 ms) observed by Federmeier et al. (2007) to unexpected words in strongly, but not weakly, constraining contexts. Our findings go a step further in showing that such positivities can be graded, based on the degree of expectancy for a particular, but not necessarily presented, continuation in a certain context. And though not reported on in their original publication, a similar effect was also observed for low cloze probability word endings in highly constraining sentences for a number of individual participants in Kutas & Hillyard's 1984 study, which manipulated sentence constraint and target noun cloze probability (from personal communication and inspection of data with M. Kutas). These unreported findings highlight a point we would like to make about why, until recently, post-N400 LPs have not been systematically analyzed and discussed in the literature. It is not the case that we believe, based on nearly three decades of N400 research, that ours and a handful of other studies are the only ones where semantically unexpected sentence continuations (or perhaps even word priming paradigms) have elicited this biphasic N400-LP pattern. For instance, Van Petten & Luka (2006) talk about a 'post-N400 positivity'; Swick, Kutas & Knight (1998) noted an absent LP following N400s to incongruent sentence endings in patients with frontal lobe damage which was present in healthy controls; Schwarz, Kutas, Butters & Paulson (1996) observed (in elderly participants) a left hemisphere LP (600-800 ms) to semantically unrelated words in a category priming paradigm; Moreno, Federmeier & Kutas (2002) found a frontal positivity (650-850 ms) in Spanish-English bilinguals in response to low cloze sentence completions as well as to unexpected code switches; and Coulson & Wu (2005) found an enhanced anterior positivity (700-900 ms) with RVF presentation to unrelated versus related probe words following joke stimuli. Researchers *have* noted LPs accompanying N400s, but as the diversity of experimental paradigms listed here attests to, the wide array of circumstances under which these effects obtain has perhaps made it difficult, or maybe undesirable, to attempt a more unifying theory of the component's functional significance. It is likely, too, that in many studies where the N400 may have been the "effect of interest", the focus may have been less on determining the functional significance of such LPs. We believe the proliferation of "semantic P600" studies in recent years - some of which have noted a similar biphasic pattern – is motivation for some of these "orphan" N400-LP studies to be reexamined.

4.13.2. Relation of current LP findings to recent "semantic P600" literature

It is also worth attempting to compare how our finding of an LP in response to constraint violations in the current study fits with some of the recent studies that have similarly found LPs to non-syntactic (semantic) experimental manipulations. In terms of sentence processing, these controversies surrounding the LP are important, because they speak to whether there is neural processing that is functionally specific to syntax or whether ERP effects like the P600/LP in fact reflect brain activity associated with more general aspects of linguistic (or perhaps not even just linguistic) comprehension. For instance, in a review article, Kuperberg (2007) summarizes and classifies some of the factors that have been shown to influence the semantic P600 (and/or attenuate the N400 effect). These include: 1) semantic-thematic fit, attraction and reparability, 2) semantic associations between verbs and arguments, 3) violation of verb thematic structure (e.g., animacy violations), 4) influence of task (e.g., plausibility judgments versus passive reading), 5) real world plausibility violations, and 6) context effects (e.g., implausible sentence continuations within supporting contexts elicit P600s, but N400s without supporting context). **Table 4.1** outlines some of these effects and their eliciting paradigms/stimuli. Our experiment finds one more piece of evidence for an LP effect that is *not* related to syntactic processing.

Another point to keep in mind is that although many of the "semantic P600" studies have similarly found positivities to experimental manipulations where N400s were originally expected, the resemblance of the LP effects across experiments is not clear. With varying latencies, time windows, scalp distributions (sometimes frontal, sometimes posterior, sometimes widely distributed) and accompanying ERP components (sometimes accompanied by an N400, sometimes not), it is quite possible that what is being discussed is not a single, unitary effect – not to mention how and if the LPs in the "semantic" studies relate to the P600s elicited in response to clearly syntactic manipulations. (Refer to **Table 4.1**). (Note, though, that Friederici, 1995, has suggested that latency variability of the P600 between studies may be more a function of the complexity of a required revision, than indexing different cognitive processing.) Although a clear picture of the functional processes that these effects are indexing is still emerging, a few research groups have offered some intriguing proposals that we consider to be more in line with the results of our current experiment.

In particular, the various collaborations of H. Kolk, D. Chwilla, and colleagues (e.g., Van Herten, Chwilla & Kolk, 2006; Kolk & Chwilla, 2007) have put forth a conflict monitoring theory of the P600 effect, suggesting that the component reflects more general executive function – reprocessing of the type that includes (but is not specific to) how syntactic violations or ambiguities are handled. Specifically, they suggest that the P600 could index "error" reprocessing at a variety of linguistic levels (e.g., phonological, lexical, conceptual, or orthographic), where a strong tendency to both accept (based on projections from preceding context) and reject (based on actual input) triggers monitoring, reprocessing and restoration. Implicit in this theory, then, seems to be an idea similar to ours – that there is indeed top-down contextual prediction (at least at some level) and that the P600 results from a clash of the representation of a pre-activated item with the representation generated by the eventual actual input. One area of potential divergence between their and our proposals is that they suggest (van de Meerendonk, Kolk, Vissers & Chwilla, 2008) that conflict monitoring triggers reanalysis, and hence P600s are present, only for highly, but not mildly, implausible events. In contrast, the graded LP effect observed to Unexpected nouns as a function of constraint-generated expectancies in our study suggests that additional processing may be triggered even when the strength of prediction is moderate and when continuations are plausible (though not anticipated). With regard to our results, we find Kolk and colleagues' proposal of the P600 as an index of more general monitoring/reprocessing a somewhat better fit than other recent explanations of the semantic LPs, which maintain that the mechanism of reprocessing is at least to some degree syntactic in nature. For instance, as a unifying theory for the variety of data in which "semantic P600s" have been observed, Kuperberg (2007) has suggested that the ERP effect arises when representations generated by two different processing streams, a more semantic one and more combinatorial (syntactic) one, come into conflict and restructuring (and possibly repair) is required. Note that the proposed combinatorial mechanism functions not only through application of morphosyntactic rules, but also through semantic-thematic constraints. However, even casting a "syntactically wide" net of what their combinatorial mechanism entails, it is difficult to conceive how our constraint violation-elicited LPs could fit with this particular type of conflict resolution. Our data show even less promise of reconciling with another proposal of the eliciting source of "semantic P600s" – that of as "semantic attraction" (Kim & Osterhout, 2005), where semantic word relations overrule the sentence grammar/thematic role assignment and lead to ungrammatical readings (e.g., *The hearty meal was devouring*).

If the LP observed in our study (as well as in other sentential N400-type studies), the recent "semantic P600" effects, and the more traditional "syntactic" P600s turn out to be the same component (or at least members of the same family of components), then indeed there will need to be even more revision to the current functional proposals. In our estimation, the commonality of the eliciting events of all three types of studies could best be described in terms of contextual pre-activation of linguistic information and the subsequent delivery of input which alters the trajectory of the (more or less strongly) predicted course. Without strong contextual constraint – i.e., strong pre-activation of upcoming semantics, syntax, phonology, or even event structure – there is no clear trajectory to be altered. But when probabilistic expectancies are high, and specific networks have been strongly pre-activated in the brain, there may be cause not only to attempt to integrate the new bit of information, but also to revise, reanalyze, repair, or inhibit the contextual representation that had been constructed up until that point. Further research will be needed in order to discern between these and possibly other potential functional explanations.

4.13.3. Outstanding questions

4.13.3.1. Inconsistency in sentence processing LP findings

One outstanding question is why some experiments with so-called "semantic" manipulations have exhibited LPs in conjunction with an N400 effect, while others have shown only a LP effect, while still others show the expected N400 but no LP effect at all? We believe that the answer lies, in good part, with the stimuli that the various experiments have used, and possibly with associated tasks. A few initial proposals for such variability are (a) that the use of congruous versus incongruous or plausible versus implausible continuations may be playing a role. While the use of anomalies has been fairly standard practice for testing and determining where the comprehension system will break down, use of clever manipulations of more naturalistic stimuli may end up being quite important for ruling out cognitive strategies that may be employed when dealing with the types of input that are rarely, if ever, encountered in everyday life. (b) Another possible source for variability may relate to the use of sentence-medial versus sentence final targets, as words in the sentence final position may be subject to additional processing that could modify how the effects appear compared to when they occur at other positions in the sentence. (c) Yet another potential factor may have to do with the overall ratio of stimulus types (e.g., congruent versus incongruent, complex versus simple, (perceived) grammatical versus ungrammatical, etc.) For instance, Coulson, King & Kutas (1998) have argued for the probability sensitivity of the P600 – larger effects when ungrammatical trials (for instance) comprise a smaller proportion of the total number of stimuli within a given block or experiment. It is not unreasonable to consider that potential P600s (or LPs) may in effect be diminished in experimental designs with larger ratios of "odd" to more "normal" sentences. (d) Along the same lines, there are indications that task demands (e.g., reading for comprehension, plausibility judgments, grammaticality judgments) may impact the elicitation or modulation of the P600, as several studies (Vissers, Chwilla & Kolk, 2007; Gunter & Friederici, 1999; Gunter, Stowe & Mulder, 1997; Hahne & Friederici, 2002) have suggested. (e) And finally, there is always the possibility of component overlap, in which N400s and LPs may in effect be canceling each other out. With variability in the latencies and scalp distributions of the "semantic P600" effects that have been reported to date, it is not clear how these two effects may be interacting.

4.13.3.2. Relation of early ERP findings to prediction

Though investigation of frontal N1 and frontal P2 effects to our manipulations of cloze probability and constraint were not the main areas of interest for the current study, our results nonetheless suggested that these components may reflect some aspects of sentence context-based prediction. In our noun cloze analysis, for instance, we found that N1 amplitude increased and P2 amplitude decreased to noun targets nouns following unexpected (relative to expected) indefinite articles. Based on previous research, these results perhaps point to the unexpected articles signaling the necessity for an increase in attentional load (in the case of the N1) and the fact that the system is less primed for visual feature extraction of an expected target when the parser has been cued that the most likely word will not (immediately) be presented. Results of from our various constraint violation analyses, though, were less consistent. Contrary to the N1 findings from our cloze analysis, we observed *increases* in N1 negativity with increasing constraint (violation) only for the EA-UN nouns – a finding which suggests a pattern of more stimulus-driven effects, rather than one guided by increasing attentional resources when the parser is less certain about an upcoming continuation. Regarding the P2, correlations using a traditional constraint violation measure indicated that as expectations for upcoming nouns increased, so did the ERP amplitude between 150-250 ms – but only for EA-UN, not UA-UN nouns; however, using an alternative constraint measure for the UA-UN nouns did reveal such a correlational pattern. These inconsistencies across the various constraint measures undoubtedly offer potential insights into the functional nature of the N1 and P2 components, though certainly any interpretation would benefit from a more systematic, and less opportunistic, manipulation of the factors which may be contributing to their amplitude modulations. Findings indicating that expectancy has its influence on such early, perception-related ERP effects would offer strong support for a more comprehensive view of contextually-guided anticipation, where "lower" level – as well as higher level – cognitive processes are tuned according to anticipated processing requirements.

4.14. Summary and conclusions

In this experiment, in addition to replicating the previous finding from Experiment 1 of prediction effects at prenominal English indefinite articles as indexed by N400 modulation to anticipated but not presented nouns, we investigated a late positive ERP deflection to semantically unexpected nouns in sentences of varying constraint. We found that this component was associated with a "cost" to violating a prediction formed by way of contextual constraint. Our observation of a graded late positivity to contextually unexpected nouns that increased as a function of how much that context constrained for a probable continuation supports a comprehension model where there is a general cost to mispredicting; in our case, this cost is reflected at a semantic level, though it is possible that the brain initiates a similar response to prediction violations along some other dimension (e.g., more syntactic). These results indicate that even though an anticipatory parser may not always "get it right", pre-activation of upcoming material could still be occurring, even when constraint is not maximal; in other words, there may be a processing cost when the parser strongly (and even not so strongly) commits to – but does not receive – an item. In addition, we note a clear double dissociation between the sensitivities indexed by the N400 and by the LP ERP components: while the N400 correlates with the probability of a particular sentence continuation regardless of the "fan" of possible likely responses, the LP is sensitive to this "fan" (or lack of "fan"), increasing in amplitude when a continuation becomes more highly preactivated, but is not received. While our study does not address the precise functional nature of the observed late positivity (whether it relates to reanalysis, inhibition, conflict monitoring, etc.), we believe that it may be similar in nature to other "semantic P600" effects which have been reported upon recently in the sentence processing literature. In particular, we propose that ours (and other such late positivities – including the P600) may not reflect syntax-specific processing at all; rather, the effect may be related to the reconciliation of predicted items with representations activated by the eventually-presented physical stimuli.

4.15. References

Altmann, G.T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280-289.

Bornkessel, I., Schlesewsky, M. & Friederici, A.D. (2002). Beyond syntax: Language-related positives reflect the revision of hierarchies. *Neuroreport: For Rapid Communication of Neuroscience Research*, 13(3), 361-364.

Bornkessel, I., Schlesewsky, M. & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes*, 18(3), 269-298.

Clark, V.P. & Hillyard, S.A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, 8(5), 387-402.

Coulson, S., King, J.W. & Kutas, M. (1998). ERPs and domain specificity: Beating a straw horse. *Language and Cognitive Processes*, 13(6), 653-672.

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

Coulson, S. & Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Research*, 1146, 128-145.

Coulson, S. & Wu, Y.C. (2005). Right hemisphere activation of joke-related information: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17(3), 494-506.

Deacon, D., Dynowska, A., Ritter, W. & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, 41(1), 60-74.

DeLong, K.A., Urbach, T.P. & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117-1121.

Engel, A.K., Fries, P. & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews. Neuroscience*, 2, 704-716.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505.

Federmeier, K.D. & Kutas, M. (1999). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Federmeier, K.D. & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 202-224.

Federmeier, K.D., Mai, H. & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition*, 33(5), 871-886.

Federmeier, K.D., Wlotko, E. W., De Ochoa-Dewald, E. & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research. Special Issue: Mysteries of Meaning*, 1146, 75-84.

Forster, K.I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. (75-107). Cambridge, MA, US: The MIT Press.

Friederici, A.D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language. Special Issue: Linguistic Representational and Processing Analyses of Agrammatism*, 50(3), 259-281.

Geyer, A., Holcomb, P., Kuperberg, G. & Perlmutter, N. (2006). Plausibility and sentence comprehension. An ERP Study. *Cognitive Neuroscience Supplement 2006*, Abstract.

Groppe, D.M. (2008). Dissertation: Common independent components of the P3b, N400, and P600 ERP components to deviant linguistic events. ProQuest Information & Learning.

Gunter, T.C. & Friederici, A.D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology*, 36(1), 126-137.

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

Hagoort, P., Brown, C. & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes. Special Issue: Event-Related Brain Potentials in the Study of Language*, 8(4), 439-483.

Hahne, A. & Friederici, A.D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13(3), 339-356.

Hillyard, S.A. & Münte, T.F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception & Psychophysics*, 36(2), 185-198.

Hoeks, J.C.J., Stowe, L.A. & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19(1), 59-73.

Kaan, E. & Swaab, T.Y. (2003). Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research*. 17(3), 621-635.

Kamide, Y., Altmann, G.T.M. & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133-156.

Kamide, Y., Scheepers, C., Altmann, G.T.M. & Crocker, M.W. (2002). Integration of syntactic and semantic information in predictive processing: Anticipatory eye-movements in German. Presented at the 14th Annual CUNY Conference on Human Sentence Processing, New York, USA, in March 2002.

Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205-225.

Kolk, H.H.J., Chwilla, D.J., van Herten, M. & Oor, P.J.W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1-36.

Kolk, H. & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, 100(3), 257-261.

Koornneef, A.W. & Van Berkum, J.J.A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54(4), 445-465.

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

Kuperberg, G.R., Caplan, D., Sitnikova, T., Eddy, M. & Holcomb, P.J. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, 21(5), 489-530.

Kuperberg, G.R., Kreher, D.A., Sitnikova, T., Caplan, D.N. & Holcomb, P.J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, 100(3), 223-237.

Kuperberg, G.R., Sitnikova, T., Caplan, D. & Holcomb, P.J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17(1), 117-129.

Kutas, M. & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161-163.

Luck, S.J. & Hillyard, S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291-308.

Mangun, G.R., Hillyard, S.A. & Luck, S. J. (1993). Electrocortical substrates of visual selective attention. In D. E. Meyer, & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*. (219-243). Cambridge, MA, US: The MIT Press.

Marslen-Wilson, W. (1989). *Lexical representation and process*. Cambridge, MA, US: The MIT Press.

Moreno, E.M., Federmeier, K.D. & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, 80(2), 188-207.

Münte, T.F., Matzke, M. & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience*, 9(3), 318-329.

Nieuwland, M.S. & Van Berkum, J.J.A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, 24(3), 691-701.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.

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

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

Osterhout, L., Holcomb, P.J. & Swinney, D.A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 786-803.

Otten, M., Nieuwland, M. & Van Berkum, J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, 8(1), 89.

Schwartz, T.J., Kutas, M., Butters, N. & Paulsen, J.S. (1996). Electrophysiological insights into the nature of the semantic deficit in Alzheimer's disease. *Neuropsychologia*, 34(8), 827-841.

Sitnikova, T., Kuperberg, G. & Holcomb, P.J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, 40(1), 160-164.

Stanovich, K.E. & West, R.F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, 7(3), 658-672.

Stanovich, K.E. & West, R.F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, 112(1), 1-36.

Staub, A. & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either...or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425-436.

Swick, D., Kutas, M. & Knight, R.T. (1998). Prefrontal lesions eliminate the LPC but do not affect the N400 during sentence reading. *Journal of Cognitive Neuroscience Supplement* (1998), 29.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

Van de Meerendonk, N., Kolk, H.H.J., Vissers, C.T.W.M. & Chwilla, D.J. (In press). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, 0(0), 1-16.

van Herten, M., Chwilla, D.J. & Kolk, H.H.J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, 18(7), 1181-1197.

van Herten, M., Kolk, H.H.J. & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2), 241-255.

Van Petten, C. & Luka, B.J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, 97(3), 279-293.

Vissers, C.T.W.M., Chwilla, D.J. & Kolk, H.H.J. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology*, 75(1), 8-18.

Vissers, C.T.W.M., Kolk, H.H.J., van de Meerendonk, N. & Chwilla, D.J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, 46(4), 967-982.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

Wicha, N.Y., Moreno, E. M. & Kutas, M. (2003a). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483-508.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003b). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wlotko, E.W. & Federmeier, K.D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001-3014.

CHAPTER 5.
EXPERIMENT 4: HEMISPHERIC INVOLVEMENT IN SEMANTIC INTEGRATION AND CONSTRAINT VIOLATION PROCESSING

5.1. Abstract

In ERP Experiment 3B we found evidence for a brain response in the form of an enhanced late positivity (LP) to low probability continuations violating highly, but not weakly, constraining sentence contexts. We argued that these results were compatible with a predictive language parser that is taxed when prediction is thwarted. This LP effect occurred in conjunction with the N400, an ERP component whose amplitude typically increases as a function of decreasing cloze probability but which is unaffected by the degree of contextual constraint violation. While N400 modulation to contextual fit has been well documented, our LP finding to constraint-based expectancy violations has much less of a precedent. In terms of anticipatory language comprehension, even less has been reported about the roles of the individual cerebral hemispheres for either effect.

Based on their ERP work using the visual half field paradigm, Federmeier and colleagues (e.g., Federmeier & Kutas, 1999; Federmeier, 2007) have proposed that with respect to language processing, the left hemisphere (LH) is more anticipatory, while the right hemisphere (RH) is more integrative. If this is the case, then we might expect to observe more pronounced LP effects for LH than RH hemisphere processing of prediction-based constraint violations. In addition, it has been proposed that unlike the LH, the RH is insensitive to message level constraint (e.g., Faust, 1998; Faust, Babkoff, & Kravets, 1995). While this has proposal has been refuted through testing using categorical variables (e.g., Coulson, Federmeier, Van Petten & Kutas, 2005; Federmeier, Mai & Kutas, 2005), hemispheric graded sensitivity to contextual semantic fit has yet to be tested in a fine-grained manner.

To this end, we conducted a lateralized ERP study that used stimulus materials similar in design to those employed by Kutas & Hillyard (1984) – a study in which sentential constraint violation and cloze probability were manipulated over their full ranges. Based on our findings from Experiment 3B and the anticipatory LH/integrative RH proposal, we expected to observe effects of constraint in the form of a post-N400 LP to less expected continuations in high, but not low, constraint contexts for the LH; conversely, we expected either no or a smaller LP effect to these same continuations presented to the RH. In general, these predictions were upheld. Over several types of analysis, our LP findings indicated that the LH more than the RH is sensitive, in a graded manner, to violations of constraint, consistent with language comprehension models that posit more top-down processing for the LH. Contrary to the RH/integrative proposal, though, we observed that when contextual constraint was high, RH processing, too, reflected a consequence to preactivating but not receiving, upcoming linguistic information

Another aim of the study was to determine whether the fully graded inverse relationship of cloze with N400 amplitude ($r \approx .9$) held for processing within the individual hemispheres. Here, our analyses revealed similarly high and widely distributed correlations of cloze with N400 amplitude over both hemispheres (on par with those for central presentation). These results signaled that both the LH and RH are sensitive to message-level contextual cues and experience similar levels of integration ease/difficulty for incoming words.

Finally, as there has been some evidence (e.g., Federmeier, Mai & Kutas, 2005; Wlotko & Federmeier, 2007; Experiment 3B of this thesis) for effects of sentential constraint on ERP components such as the frontal N1 and P2, we investigated but found no evidence for lateralized effects of constraint violation and cloze probability over these time windows.

In conclusion, we will attempt to relate this set of findings to a more comprehensive model of hemispheric distinctions in language comprehension.

5.2. Introduction

The brain's two cerebral hemispheres differ along a variety of anatomical dimensions. For instance, in most right-handers, the planum temporale – a brain region that is part of Wernicke's area – is larger in the left hemisphere (LH) than the right hemisphere (RH). Researchers have also proposed more functionally distinct columnar systems per surface unit in LH Brodmann's area 22 (Wernicke's area) than in the RH homologue (Galuske, Schlote, Bratzke & Singer, 2000). Others have observed more extensive higher-order dendritic branching in LH speech areas (e.g., Broca's area) than in RH homologues (Scheibel, Paul, Fried, Forsythe, Tomiyasu, Wechsler, Kao & Slotnick, 1985). Additionally, neurochemical asymmetries, which some have suggested may have been conducive to the LH organizing itself for speech in the first place, have also been described (Tucker & Williamson, 1984). However, linking such anatomical asymmetries to actual brain functions (and in particular, to behaviors related to language processing) has, over the years, proven a challenging enterprise, to say the least.

The hemispheres also vary along a number of functional dimensions, the most striking of which is the LH's specialization for language. Dating back to the mid 1800s and Paul Broca's seminal findings, the LH has been considered the "seat" of language, showing devastating effects for language processing when certain regions of this hemisphere are damaged. Over the years, however, studies using a variety of methodologies have revealed that the right hemisphere (RH), too, is capable of processing language, albeit with a different set of specializations than the left hemisphere. For instance, we know from lesion data that the RH is associated with linguistic activities such as the processing of non-literal language and prosodic contour, as well as the ability to extract the gist or theme from a discourse. Less well understood – and of particular interest for the current study – though, is how or the degree to which the hemispheres may differ in their abilities to extract precise, message-level meaning from a sentence or discourse and the (possibly variable) mechanism(s) under which they may be operating. But first, in order to investigate the roles of the hemispheres in language processing – and in particular anticipatory language processing – it may be informative to keep in mind research that has examined cognitively "broader" functional asymmetries for the brain's two halves.

5.2.1. Some proposed general functional asymmetries

Wolford, Miller & Gazzaniga (2000), for instance, used a probability-guessing paradigm with split-brain patients to test the problem-solving strategies employed by the two hemispheres. When healthy normal individuals were randomly presented with frequent (e.g., 75%) and infrequent (e.g., 25%) stimuli and asked to guess about the nature of the upcoming stimulus before each trial, participants typically adopted a "matching" strategy, where they attempted to pattern their percentage of guesses on the frequency pattern of the stimuli. This contrasts with the strategy employed by most other animals, who tend to "maximize" by continually guessing the most frequent stimulus. The non-human animals, it turns out, do better than their human counterparts – a result which the researchers attribute to humans' propensity for pattern extraction. Interestingly, when the same series of stimuli are presented separately to the hemispheres of split-brain patients using a visual half-field paradigm, the LH more consistently employs the matching strategy, where the RH is more prone to the maximizing strategy. Though non-adaptive, the researchers propose that the LH "interpreter" cannot help but hypothesize some kind of pattern, even when there is none. Gazzaniga and colleagues find additional support for this LH interpreter based on their many observations of split-brain patients, similar to the following one (Gazzaniga, 1985). A split-brain patient is presented with a picture shown exclusively to the left hemisphere (e.g., a chicken claw) and one shown exclusively to the right hemisphere (e.g., a snow scene). Afterwards the patient is presented with an array of pictures and asked to choose the ones associated with to the lateralized pictures. While he correctly chooses pictures of a shovel for the snow scene with the left hand and a chicken for the claw with the right hand, he interestingly justifies these choices by saying that "the chicken claw goes with the chicken, and you need a shovel to clean the chicken shed." The researchers argue that findings like this offer evidence that the left brain tries to interpret the actions of both hemispheres. The flip side of a confabulating left hemisphere, Metcalfe, Funnell & Gazzaniga (1995) propose, is a more veridical right hemisphere, which they suggest is evidenced by memory studies in which the right hemisphere more correctly identifies previously viewed items and correctly rejects new items. In contrast, the left hemisphere is more prone to false identification of new items when they are similar to old items.

At other levels, hemispheric differences could be considered in terms of attentional biases, with the LH better at attending to local features in visual-spatial tasks, and the RH better a directing attention to global aspects (Yamaguchi, Yamagata & Kobayashi, 2000). These differences could affect how information is encoded and in turn retrieved from memory during language processing. And in terms of memory, numerous studies have suggested that the LH preferentially processes verbal information (e.g., letter strings or words) while the RH is biased toward non-verbal material (e.g., faces, line orientations). Other theories suggest that the LH is involved in retrieval of information from semantic memory and the RH in retrieval from episodic memory (Tulving, Kapur, Markowitsch, Craik, Habib & Houle, 1994).

Undoubtedly, the story for functional hemispheric asymmetries is a far richer one than can be explored here, the main point being that it would make sense that a higher order process such as language comprehension could be driven by hemispheric effects at broader levels of cognition - especially if one takes the view that language processing does not employ a "special language processor", but rather is a "new machine built out of old parts" (Bates, 1988). These are considerations worth keeping in mind as we investigate the role of the hemispheres within the domain of language processing, and specifically at the level of sentence context effects and violations of expectancy.

5.2.2. Word priming studies

The bulk of existing work exploring the nature of hemispheric asymmetries in semantic language comprehension has been conducted at the lexical, rather than at the sentence or discourse, level, using word priming paradigms. Many of these studies have relied on the visual hemi-field technique, in which stimuli are presented a few degrees to the right or left of fixation in order to expose only the contralateral hemisphere to that stimulus for the first approximately 10 ms (Banich, 2003). The consequence of this short head start in apprehending the stimulus results in hemispheric processing differences that carry over even into relatively late stages of processing, which, by inference, reflect how the two hemispheres deal with semantic relationships and sentential constraint, among other things. From a variety of hemi-field word priming experiments, a few main themes have emerged. Under one theory, the right hemisphere processes semantic information at a much "coarser" level than the left hemisphere, tending to weakly activate multiple meanings, rather than selecting for a single relevant meaning (Beeman, 1998; Jung-Beeman 2005). This idea is supported by Beeman's work (e.g., Beeman, Friedman, Grafman, Perez, Diamond, & Lindsay, 1994) using summation primes (i.e., three prime words loosely related to a given target). In these studies, when a word trio such as "foot-cry-glass" precedes a target word like "cut", the RH exhibits a greater priming effect, suggesting that it is more sensitive to distant relations. Findings by Chiarello and colleagues (e.g., Chiarello, 1998) argue along the same lines, with both the LH and RH showing priming for strongly related category members (e.g., “arm-leg”) but only the RH showing effects for distantly related category pairs (e.g., “arm-nose”).

Another theme that has emerged from the priming literature is that the LH processes language faster than RH. For instance, in a version of the “arm-nose” word priming experiment described above where stimulus onset asynchrony (SOA) was manipulated, Koivisto (1997) showed that distantly related pairs could also show priming in the LH, but only at short SOAs (165 ms). At longer SOAs (e.g., 500 and 750 ms), priming effects were restricted to the RH. Burgess & Simpson (1988) note a similar SOA/hemispheric pattern in a word priming study employing subordinate meanings of ambiguous words (e.g., “bank-river”). The takeaway message here seems to be that the LH may initially activate more distantly related items, but quickly hones in on primary, more probable meanings.

Still another proposal of asymmetries in language processing focuses on different systems of semantic coding (rather than size of semantic fields) in the two hemispheres. For instance, Deacon, Grose-Fifer, Yang, Stanick, Hewitt & Dynowska (2004) – using a semantic word priming study – have recently suggested that the LH is more sensitive to holistic and associative (e.g., “honey-bee”) rather than featural or categorical relationships between items, which allows for priming via spreading activation: the RH, on the other hand, represents semantic content more on the basis of semantic features, with priming arising when there is featural overlap (e.g., “broccoli-tree”) between items.

5.2.3. Sentential level studies

To a lesser extent, experimental work has addressed hemispheric processing differences at the levels of sentence and discourse comprehension. One proposal (Chiarello, 2000; Faust, 1998; Faust & Kravetz, 1998) posits that while the LH is capable of integrating information at various linguistic levels to form message-level representations, the RH constructs meaning more on the basis of word-level association, in a bottom-up fashion. This proposition stems from behavioral studies manipulating sentence constraint, in which word continuations presented to the LH showed graded facilitation as indexed by lexical decision times: the RH, on the other hand, benefitted only from the highest levels of constraint. In another study (Faust, Babkoff & Kravetz, 1995), scrambled sentential word order led to priming effects similar to those for congruent sentences in the RH, whereas the LH benefitted only from properly ordered sentences.

In contrast to this RH "message-blind" model, Coulson et al. (2005) combined the visual hemi-field paradigm with event-related brain potentials (ERPs) to pit the effects of word-level versus sentential-level priming. Using ERPs in conjunction with hemi-field studies offers a methodological advantage over behavioral tasks such as lexical decision task or word naming, in that potential hemispheric asymmetries related to producing task-related outputs can be avoided. Presenting associated word pairs in isolation, they noted similar ERP priming effects over both hemispheres. When the same words pairs were embedded in sentences, both hemispheres yielded similar context effects, as indexed by reduced N400s to congruous endings. Decreases in N400 amplitude are thought to reflect the degree to which an item semantically "fits" within its surrounding context. These findings thus challenge the models of Faust and colleagues and argue for RH language processing that is not insensitive to message-level constraint.

Yet another explanation for hemispheric asymmetries in sentence processing – and one which is highly relevant for this thesis – has been offered by Federmeier and colleagues (Federmeier & Kutas, 1999 and 2002; Federmeier, 2007). They also used online measures afforded by the ERP methodology to demonstrate that sentence level constraint may help facilitate semantic language comprehension for both hemispheres, but in somewhat different ways. Based on their hemi-field studies, they argued that while right visual field

(RVF)/LH language processing may be oriented more toward prediction and the use of top-down cues, the left visual field (LVF)/RH processing may be more biased toward the integration of linguistic information with working memory. In their 1999 study, the researchers compared lateralized expected sentence completions to within and between category violations in high and low constraint sentences. For high constraint contexts such as '*He caught the pass and scored another touchdown. There was nothing he enjoyed more than a good game of…*' there were three possible continuations: the expected item (*football*), an unexpected item from within the same category (*baseball*), or an unexpected item from a different category (*monopoly*). Although both violation types showed larger N400s than that to the expected item (*football*), the N400 to the within category violation (*baseball*) was reduced relative to that of the between category violation (*monopoly*) for RVF(LH) presentation only. Although within and between category violations were judged implausible for both high and low constraint sentences, the RVF(LH) within category violation N400 was reliably reduced only when target nouns were highly constrained. The researchers suggested that this pattern of results in the LH could only be explained by the greater overlap in perceptual and semantic features of the expected exemplar with the within category violation compared to the between category violation. They proposed that in high constraint sentences, contextual information had already acted via semantic memory to preactivate some of the features of the expected exemplar. In contrast, they proposed that the lack of facilitation for within category violations in high constraint contexts for LVF presentation suggested that the RH acted in a more a bottom-up fashion, integrating input only at the point at which it was received.

These results from Federmeier & Kutas' study are difficult to reconcile with some of the alternative theories of hemispheric semantic processing mentioned above. First, taking Beeman's coarse coding hypothesis, the *football/baseball/monopoly* data do not seem to support the concept of a RH with less focused activations compared to the LH, because, in fact, the RH distinguished (to an even greater degree than the LH) between highly expected sentence endings and related items. This argues against the RH processing context information in an inexplicit, broad manner. As for the Deacon et al. (2004) proposal that semantic information is activated on the basis of featural overlap in the RH but more on associative links in the LH, this explanation is not compatible with the N400 facilitation observed for *baseball* in the LH but not RH in Federmeier & Kutas' study: the relationship between *baseball* and the expected ending (*football*) primed by the sentence context is one based more on shared features rather than association. And finally, proposals like those of Chiarello (2000) and Faust (1998), who suggest that the RH, but not LH, is "message blind", do not explain the pattern of N400 facilitation for *baseball* in the LH, relative to an equally implausible continuation like *monopoly*. Thus, Federmeier & Kutas' explanation of their data patterns in terms of predictive (LH) and integrative (RH) processing, at least for now, seems like a promising proposal.

In Experiments 1, 2 and 3B of this thesis, we observed an enhanced late positivity following the N400 that appeared to increase in amplitude as a function of how strongly sentential constraint was violated. Linking these findings with a potential "cost" to mispredicting, our results suggested that on the whole, the brain engages in predictive processing and is sensitive to not receiving highly expected linguistic input. With Federmeier & Kutas' LH predictive/RH integrative proposal in mind then, we wondered whether the LP effect observed in our studies might turn out to be larger for targets processed by the left, compared to the right, cerebral hemisphere. If this were the case, it would be compatible with the argument for a more predictive LH. More generally, with two cerebral hemispheres, and research suggesting that these two halves may contribute non-identically to construction of meaning based on sentence constraint cues, we wanted to investigate their sensitivities in processing contexts of varying constraint and continuations that ranged in expectancy.

These questions have been preliminarily addressed by Wlotko & Federmeier (2007), who conducted a lateralized version of a study by Federmeier, Wlotko, De Ochoa & Kutas (2007) – a central presentation ERP experiment for which they observed an LP (500-900 ms) similar to ours for low relative to high cloze continuations in high constraint ('*He bought her a pearl necklace for her…birthday/collection*') but not low constraint ('*He looked worried because he might have broken his…arm/collection*') contexts. With their LP finding in central presentation, they argued (similar to our Experiment 3B) for a cost – perhaps reflecting inhibition or revision – when processing unexpected words in highly predictive contexts. However, for lateralized presentation of the same stimuli, they observed that the two VFs not only produced different, atypical (amplitude not strictly graded by cloze probability) N400 patterns, compared to effects for those same stimuli during centralized presentation, but that the LP effect present for centralized presentation was unexpectedly absent for both visual fields (VF) of presentation. These results were not anticipated based on the predictive LH/integrative RH account. The authors argue that the N400 and LP effects present only during centralized presentation may reflect processing that emerges only through cooperation of the two hemispheres. Unspecified by the authors, though, is why the brain's language processor would show a LH bias for prediction (as Federmeier & Kutas argued for in their 1999 and 2002 studies), but not for the consequences of mispredicting; in other words, if the LP is related to prediction, it would have been expected that the effect might have been stronger (or at least present) in the LH compared to the RH. The authors state,

> Because this was the first visual half-field study to include plausible completions with several levels of cloze probability, further work manipulating cloze probability in a fine-grained manner will be needed to fully ascertain how the response functions of two hemispheres are

influenced by off-line expectancy and how the pattern observed with non-lateralized presentation relates to that seen from each VF individually.

5.2.4. The current study

We attempted to address these issues by undertaking the current study. Though our main objective was to investigate potential hemispheric differences in semantic sentence comprehension, our aims within this overarching goal were twofold. First, we wanted to test, using sentence/lateralized noun target stimuli with a full range of contextual constraint/cloze probability, whether there was a hemispheric bias in the processing reflected by the LP observed in Experiments 1, 2 and 3B. If, indeed, LH sentence comprehension is more anticipatory and RH comprehension more integrative, we might expect enhanced LP effects for RVF, but not LVF, presentation. Specifically, we propose that by comparing effects of constraint on the processing of low cloze sentence continuations (thus holding cloze constant), LP amplitude would be modulated in a systematic, and perhaps differential, way for one or both hemispheres.

A second goal of the current study was to capitalize on the extended range of constraint and cloze probability in our stimulus set to systematically examine potential hemispheric influences of constraint and cloze on N400 amplitude. While a number of behavioral studies using priming measures and lexical decision tasks have postulated that the RH comprehension occurs in a “message-blind” manner (e.g., Faust, Kravetz & Babkoff, 1993; Faust et al., 1995; Faust & Kravetz, 1998), several ERP studies have demonstrated that N400 congruity effects as well as effects indicating sensitivity to degree of message level constraint can be obtained within both hemispheres (e.g., Coulson et al., 2005; Federmeier & Kutas, 1999; Federmeier et al., 2005). However, the correlational relationship of N400 mean amplitude across a range of cloze probability for central presentation (as demonstrated by Kutas & Hillyard, 1984, and in the previous ERP studies of this thesis) has never been tested with visual hemi-field presentation. Kutas & Hillyard's landmark ERP study used sentence contexts with three levels of constraint (high, medium, and low) and two or three levels of cloze within each constraint level (high, [medium], low). Note that "high" cloze continuations for "high" versus "medium" or "low" constraint contexts are, by default, parameterized differently; for instance, a "high" cloze continuation for a low cloze sentence may only have a cloze rating of 15%, whereas a "high" cloze continuation for a medium constraint sentence might have a 60% cloze rating. Importantly, though, "low" cloze continuations for each level of constraint exhibit similarly low cloze values. With similar low cloze averages across constraint level, the factor of cloze can be held constant while constraint is manipulated. One of several findings demonstrated in this study was that N400 amplitude, which turned out to be highly inversely correlated with an item's cloze probability ($r \approx -.9$), is insensitive to the degree to which a context's constraint is violated; in other words, regardless of the cloze probability of the most expected continuation for a particular context, N400 amplitude is consistently large when a low probability item is provided instead. Or as the authors put it, "The N400 reflects the extent to which a word is semantically primed, rather than its being a specific response to contextual violations." Yet another way of putting it is that the N400 does not appear to reflect a "cost" for unfulfilled prediction.

In part, it has been the N400's lack of sensitivity to constraint violation that has made difficult the task of determining whether language comprehension occurs in a more anticipatory than integrative fashion. In the most straightforward of "N400" experimental designs (e.g., high versus low cloze sentence continuations), the component/effect cannot differentiate between the two theories, at least at the point when the target word is presented. (Note, though, how the careful experimental manipulations like those of Federmeier & Kutas, 1999a,b, yielded results at the target items that allowed them to argue in favor of prediction.) However, as Experiment 3B of this thesis suggests, there may be some evidence for the consequence of failed prediction – evidence that has over the years remained elusive and whose absence has likely been an obstacle to wider acceptance of anticipatory comprehension models. Though the N400 component is insensitive to constraint violation, the late, post-N400 positivity (the LP) we observed with centralized presentation appears to increase in amplitude to low cloze continuations the more highly constraining the sentence context is. These results in Experiment 3B were evident even though we opportunistically relied on post-hoc analyses of the partial ranges of constraint and cloze within this experiment. A logical follow-up, then, for verifying the influence of constraint violation on LP amplitude would be to construct experimental stimuli with a full and systematic range of constraint and cloze probability (the same type of stimuli used by Kutas & Hillyard, 1984). One may indeed wonder why LP effects were not, then, observed in this study dating back a quarter of a century. Actually, though not reported in their original publication, a similar effect was observed for low cloze endings in strong relative to weak constraint sentences for a number of individual participants (per personal communication with M. Kutas) and is even suggested by visual inspection of the averaged waveforms in Figure 1B of the original publication.

Thus, in the current study, we aim to dissociate the effects of cloze probability and constraint violation (as operationalized by cloze probability values from the sentence context norming), and by doing so, examine potential hemispheric indicators of semantic integration and expectancy violation. We will utilize the existing sentence stimuli from Experiments 1 and 3B of this thesis, supplementing these lists with items that round out the <1-100% range of constraint and cloze conditions. Language contexts can range from being highly constraining (maximum constraint would mean that a context would have a single, plausible continuation) to weakly constraining (a context could plausibly continue in a potentially unlimited number of different ways). In the case of a high constraint context, the constraint value (e.g., 95% constraint) is determined by and thus has the same value as the highest cloze item (95% cloze). The same process determines the constraint value of a low constraint context (e.g., a context for which the most common continuation is provided by only 12% of respondents also is said to have 12% constraint). While factors of constraint and cloze probability are confounded at their upper ends, they are dissociable for sentences which do not continue with the highest cloze item for that context (e.g., a context with 95% constraint can continue with a word with only 5% cloze). The comparison of low cloze items in high constraint (e.g., '*He mailed the letter without a…thought*') versus low constraint ('*He was interrupted by a…thought*') contexts is a potentially revealing one in terms of anticipatory processing, because with low cloze probability held constant, one can more clearly examine the contribution of the factors leading up to, as well as factors triggered upon presentation of, the target item.

By using full ranges of constraint (3 levels: high, medium, and low) and cloze probability (2-3 levels: high, [medium], and low) similar to those used by Kutas & Hillyard (1984), we hope that this fine-grained manipulation of experimental factors will help us to draw more informed comparisons about the contributions of LH versus RH sentence comprehension. (See **Table 5.1** for sample stimuli.) Our analyses will focus on contrasting N400 and LP effects across the two VFs of presentation using statistical comparisons (ANOVAs as well as correlations) that will allow for examination of potential effects at different levels of granularity. In addition to factors of cloze probability and constraint often being conflated in previous studies, studies which have attempted to dissociate the two variables most frequently have examined the factors in a binary fashion – i.e., high/low levels of cloze or constraint. We believe our range of constraint and cloze values, which make correlational analyses possible, may offer even deeper insights into the nature of these effects.

Table 5.1. Sample experimental sentence stimuli.

Constraint	Cloze	Sample Stimuli	Constraint	Cloze
HIGH	high	*Bart did not clean his wound properly. He ended up getting an **infection** soon after.*	97%	97%
	low	*For the snowman's eyes the kids used two pieces of coal. For his nose they used a **berry** from the fridge.*	100%	<1%
MEDIUM	high	*The pilot had to make an emergency landing in the middle of the desert. The plane was nowhere near an **airport** or a road.*	78%	78%
	medium	*Joe went to the hardware store. He bought a **hammer** for half price.*	55%	55%
	low	*The cat climbed up the bird feeder. When he reached the top he saw a **squirrel** and pounced on it.*	56%	6%
LOW	high	*Andy was looking for a place to plug in the lamp. He searched high and low but couldn't find a **socket** anywhere.*	19%	16%
	low	*The backpacker was hiking through the forest. She reached a clearing and spotted an **elk** and began to run.*	16%	3%

5.2.5. Analyses and possible outcomes

Based on findings from the centralized ERP Experiments in this thesis, as well as those from Federmeier, Wlotko, de Ochoa & Kutas (2007), we anticipated that low cloze noun continuations in the current study would elicit an increased LP in high, relative to low, constraint contexts. In particular, under the predictive LH/integrative RH theory, we project that any effects of LP as a function of constraint violation will be greater (or present only) for RVF(LH) stimuli. However, as Wlotko and Federmeier (2007) did not observe lateralized LP effects to constraint violations in an experimental paradigm similar to ours, we must keep these results in mind. It is possible though, that with the full range of constraint and cloze in our stimuli that the LP may be revealed in our studies (perhaps through correlational analyses), where it was not in theirs. Observing an LP to constraint violations that is stronger for LH than RH processing would reinforce not only the idea of linguistic preactivation in general, but would offer additional support for a distinction in the way the two hemispheres use sentence context during semantic language comprehension.

Additionally, to investigate the LP, we plan to utilize an alternative analysis of contextual constraint, similar to that performed in Experiment 3B. In the norming data from that study we observed that for a truncated sentence such as, '*Because they were playing baseball so close to the house, the children ended up shattering an…*', the presentation of the unexpected indefinite article *an* drastically altered the "constraint" value of the context from when it was normed with the expected article *a*. With *a*, there was a single, highly expected continuation (*window*) with 96% cloze probability, which by definition means that the context's constraint value is also 96%. With the opposite article (*an*), the fan of responses was much wider (e.g., [*old, expensive, oval, elegant, open, upstairs, antique*] *window*): calculating cloze probability/constraint based on the most common response – in this example, *old window* – the constraint was reduced to 27%. However, if constraint is considered in a different way (e.g., the degree to which norming respondents salvaged the most expected noun by means of any intermediary adjective) the constraint value for this particular example only experiences a slight reduction, to 92% cloze/constraint (for [any adjective] + *window*). Being able to contrast ERP results for constraint operationalized in two different ways could ultimately be useful for determining the sensitivities of the hemispheres to expectancy violations.

The fact that the target nouns in our stimuli are always preceded by either *a* or *an*, also allows us to examine the role that the indefinite articles may play in shaping contextual constraint and cloze probabilities for upcoming items. Because an indefinite article has the potential to, for instance, downgrade the constraint rating of a context from high (e.g., 90%) to low (e.g., 10%), we additionally performed an analysis comparing low constraint contexts that were classified as such because of such a decrease, versus those low constraint contexts that were deemed such simply because the overall semantic information did not narrow the scope of possible continuations.

Although the N400 patterns observed for lateralized noun targets in this experiment for the individual VFs (hemispheres) cannot speak to the prediction question, they can inform us about potential hemispheric similarities (or dissociations) in semantic processing. While we have no reason to anticipate hemispheric differences on the N400 as a function of cloze probability, again we must keep in mind the findings from Wlotko & Federmeier (2007), which, using a roughly similar experimental paradigm, indicated that lateralized N400s did not pattern directly with the cloze probability for either hemisphere. Finally, based on central presentation findings (Kutas & Hillyard, 1984), we would not expect different levels of constraint (with cloze probability held constant) to modulate N400 mean amplitude, as the degree of constraint violation has repeatedly been shown not to influence this component.

In addition to investigating the N400 and LP as our primary ERP components of interest, we will also examine the influence of constraint and cloze probability on two early ERP components – the frontal N1 and P2 – as we did in Experiment 3B of this thesis. In that study, these components were seen to decrease and increase (respectively) in amplitude for noun targets continuing high constraint relative to low constraint contexts. A similar pattern of P2 results, in particular, has been noted in previous sentence processing studies, primarily for lateralized presentation (Wlotko & Federmeier, 2007; Federmeier, Mai & Kutas, 2005; Federmeier & Kutas, 2002.) Federmeier and colleagues in fact argue that the P2 effects they observed (for RVF, but not LVF presentation) are compatible with the idea of a more predictive LH.

In sum, if the LH indeed engages in more anticipatory semantic comprehension than the RH by utilizing contextual constraints to preactivate linguistic material, we expect to observe an enhanced LP to constraint violations relative to expected sentence continuations for RVF presentation. Conversely, if there is no difference in hemispheric processing on the LP, this argues against a clear predictive/integrative processing distinction between the hemispheres. We also anticipate graded effects of cloze probability on N400 amplitude for both hemispheres, but no modulation of this component as a function of constraint violation. Finally, if the effects of preactivation extend to tuning of attentional resources or more perceptual, form-based processing, we might expect these to influence early ERP effects such as the frontal N1 and P2.

5.3. Methods

5.3.1. Materials

5.3.1.1. Sentence stimuli

We used sentence stimuli that encompassed a range of contextual constraint and had target continuations with a range of cloze probability. Each stimulus item was comprised of two sentences: a context sentence and one containing a plausible, sentence-medial, left or right lateralized target noun. In order to utilize stimulus materials from previous studies, our sentence targets were always preceded by an indefinite article, *a* or *an*, though for the present experiment, we did not analyze ERPs to these words. There were 240 different sentence context pairs, each of which had two different target continuations – a consonant-initial noun or a vowel-initial noun. See **Table 5.2**.

Table 5.2. Number of stimuli for Experiment 4.

Sentence pair stimuli	2 target continuation types	2 visual fields of presentation	Total number of stimuli
240 items	consonant-initial noun (preceded by *a*)	LVF	960
		RVF	
	vowel-initial noun (preceded by *an*)	LVF	
		RVF	

The resultant 480 items, which across subjects were presented to the two visual fields (LVF and RVF, yielding a total of 960 lateralized stimuli), were then classified into three different levels of constraint (HIGH, MEDIUM, or LOW), as a function of the cloze probability of the most frequently produced item when the contexts were normed with the indefinite articles provided. The following cloze probability criteria were used to parameterize the three levels of contextual constraint: HIGH (HI) constraint = 80-100% cloze of the most commonly supplied word, MEDIUM (MD) constraint = 20-80% cloze of the most commonly supplied word, LOW (LO) constraint = 0-20% cloze of the most commonly supplied word. Within each level of constraint, noun targets ranged in cloze probability: high (hi)/low (lo) cloze for the HI and LO constraint conditions and high/medium (md)/low cloze for the MD constraint condition. See Table 5.3.

Table 5.3. Constraint/cloze breakdowns of stimuli for Experiment 4.

Constraint Level	Constraint Range	Cloze Level	Cloze Level Cloze Range Parameters	Constraint/ Cloze Condition Bin	Mean Bin Cloze	Mean Cloze Per Constraint Level	No. of Items per Bin
HIGH (HI)	80-100%	high	80-100%	HI/hi	92%	66%	100
		low	0-16%	HI/lo	4%		41
MEDIUM (MD)	20-80%	high	60-80%	MD/hi	70%	32%	50
		medium	21-59%	MD/md	36%		92
		low	0-20%	MD/lo	5%		87
LOW (LO)	0-20%	high	6-19%	LO/hi	11%	7%	61
		low	0-5%	LO/lo	2%		49
							480 **Total**

Due to the multiple levels of constraint and cloze probability, this was not a crossed design; however, the two stimulus lists (X 2 visual fields of presentation, which yielded 4 lists total) were matched on factors of word frequency (Kucera & Francis written frequency count, 1967), word length, orthographic neighbors (Medler & Binder, 2005, MCWord database), concreteness (MRC Psycholinguistic Database), and overall cloze probability (**Table 5.4**). Each list contained 240 sentence pairs, with equal numbers of *a* and *an* continuations. Only one version (*a* or *an* continuation) of a particular context was included per list. **Table 5.1** shows examples of sentence contexts and targets for each level of constraint/cloze.

Table 5.4. Stimulus factors controlled for in Experiment 4, with means and standard error of the mean (SEM) information.

List	Statistic	Cloze probability	Word length	Orthographic neighborhood	Concreteness rating	Word frequency
A	**Mean**	0.37	6.74	2.38	296	49.79
	SEM	0.02	0.15	0.25	17.11	5.82
B	**Mean**	0.36	6.45	2.41	307	54.26
	SEM	0.02	0.14	0.27	16.91	8.73

5.3.1.2. Comprehension questions

Approximately one half (120) of the sentences in each list (240) were followed by yes/no comprehension questions pertaining to the immediately preceding sentence pair. For two of the lists, 44% of the questions followed LVF targets, 56 % following RVF targets: the opposite proportions held for the other two lists. Across all four stimulus lists, an equal number of questions followed right and left lateralized stimuli.

5.3.1.3. Cloze procedure

Cloze probabilities were obtained for the 240 experimental sentence contexts, norming each context in two different ways: 1) with the *a* indefinite article provided, 2) with the *an* indefinite article provided. Each variation was normed in an off-line sentence completion task by 31-32 University of California, San Diego (UCSD) student volunteers, who were compensated either with experimental credit or cash. Participants were instructed to continue each context with the word(s) they felt best completed each sentence. It should be noted that the indefinite articles, per se, are not relevant for the purposes of the current experiment; the stimulus set was in part derived from items used in previous studies, including Experiments 1 and 3B. Cloze probability for a given noun in a particular context was calculated as the proportion of individuals continuing a particular context with that particular word. Constraint values of the contexts were determined using the norming values for the most expected item in a particular context.

5.3.2. ERP Participants

Thirty-two volunteers UCSD undergraduates participated in the experiment for course credit or for cash. Participants (23 women, 9 men) were all right-handed, native English speakers with normal or corrected-to-normal vision, ranging in age from 18-24 years, with a mean age of 19.8 years. Eleven of the 32 participants reported a left-handed parent or sibling. Two additional participants were excluded from the analysis due to excessive eye blinks or movements.

5.3.3. Experimental procedure

Volunteers were tested in a single experimental session conducted in a soundproof, electrically-shielded chamber. They were seated in a comfortable chair one meter in front of a computer monitor and were instructed to read the sentence pairs for comprehension. They were told that for each RSVP sentence, one word would be presented to the left or right of a fixation dot. They were also informed that some sentences would be followed by yes/no comprehension questions, to which they should respond by pressing one of two hand-held buttons. Response hand was counterbalanced across participants and lists. There was a brief practice session that contained a variety of constraint conditions and targets presented to both the left and right hemi-fields, as well as comprehension questions following each item. During the practice session, eye movements were carefully monitored by the experimenter, and feedback was given to the participants, to ensure that the lateralized targets were being viewed without horizontal eye movements. For the practice session only, participants reported whether or not they had seen the lateralized noun target, and if so they were asked to verbally report the item that they read. By the end of the practice session (8 items), we were assured that all participants were able to read the lateralized words peripherally, with their eyes fixated on a central point. Participants were asked to remain still during testing, and to avoid blinking and moving their eyes while the sentences were being presented. Stimuli were presented in 10 blocks of 24 items each. Participants were given a short break after each block.

Sentence stimuli were presented visually in black type on a white background on a cathode-ray tube screen. Each stimulus item was comprised of a context sentence, presented in its entirety on screen, which participants advanced from (self-paced) with a button press. The second sentence, containing the lateralized target item, was presented RSVP. Each RSVP sentence began with a centrally presented series of crosses, on screen for a duration jittered between 1000 and 1500 ms, to orient participants to the center of the screen. The words in the RSVP sentences were centrally presented (except for the target), for a duration of 200 ms and a stimulus onset asynchrony of 500 ms. Lateralized noun targets were presented randomly to the left or right visual hemifield, with the inner edge 2° visual angle from fixation. A fixation dot remained on screen throughout the trial, positioned 0.5° below the central text baseline. Participants were instructed to remain focused on this dot throughout the sentence. 2.5-3.5 seconds of blank screen followed the offset of the sentence final word, after which a comprehension question either did or did not appear in full on the screen. If there was a comprehension question, the participant's button-press response served to advance the screen to the next sentence. Following questions, there was one second of blank screen prior to the next sentence automatically appearing on screen.

5.3.4. Electroencephalographic recording parameters

The electroencephalogram (EEG) was recorded from 26 electrodes arranged geodesically in an electro-cap, each referenced on-line to the left mastoid. These sites included midline prefrontal (MiPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial frontal (LMFr and RDFr), left and right lateral frontal (LLFr and RLFr), midline central (MiCe), left and right medial central (LMCe and RMCe), left and right mediolateral central (LDCe and RDCe), midline parietal (MiPa), left and right mediolateral parietal (LDPa and RDPa), left and right lateral temporal (LLTe and RLTe), midline occipital (MiOc), left and right medial occipital (LMOc and RMOc), and left and right lateral occipital (LLOc and RLOc). See **Figure 5.1**.

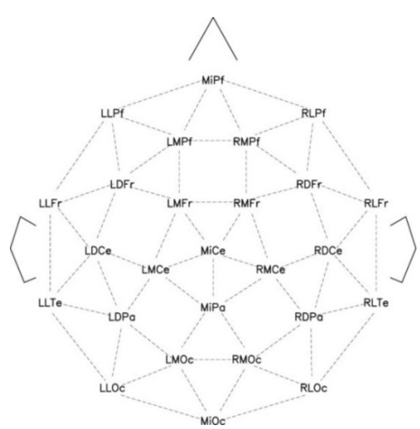


Figure 5.1. Schematic showing the array of the 26 scalp electrodes from which ERPs were recorded.

Blinks and eye movements were monitored from electrodes placed on the outer canthi of each eye and under each eye, each referenced to the left mastoid. Electrode impedances were kept below 5 KΩ. The EEG was amplified with Grass amplifiers set at a band-pass of 0.01 to 100 Hz, and continuously digitized at a sampling rate of 250 Hz.

To follow-up on interactions with electrode site, all distributional analyses reported on were conducted via an ANOVA which included the following factors: two levels of hemisphere (left vs. right), two levels of laterality (lateral vs. medial), and four levels of anteriority/posteriority (prefrontal vs. frontal vs. parietal vs. occipital). This and all subsequent distributional analyses used 16 electrode sites divided into left lateral sites (from front to back: LLPf, LLFr, LLTe, LLOc), left medial sites (LMPf, LMFr, LMCe, LMOc), right medial sites (RMPf, RMFr, RMCe, RMOc), and right lateral sites (RLPf, RLFr, RLTe, RLOc).

5.3.5. Data analysis

Trials contaminated by eye movements, excessive muscle activity, or amplifier blocking were rejected off-line before averaging. On average, 13% of LVF and 15% of RVF trials were eliminated for data time-locked to the target nouns. ERPs were computed for epochs extending from 500 ms pre- to 1540 ms post-stimulus onset. Data with excessive blinks were corrected using a spatial filter algorithm. A digital band-pass filter set from 0.2 to 15 Hz was used on all data to reduce high frequency noise. The data were re-referenced off-line to the algebraic sum of the left and right mastoids and averaged for each experimental condition, time-locked to the onset of the target nouns.

5.4. Behavioral results

Comprehension accuracy was calculated for the yes/no probe questions. Participants correctly answered an average of 96.0% (median = 96.2%, range = 91.7% to 99.2%) of the questions, indicating that they were attending to and comprehending the experimental sentences during the recording session. Overall, 57% of incorrect responses followed stimuli with LVF presentation, while 43% of incorrect responses followed RVF stimuli. Behavioral accuracy for questions following stimuli presented to the two VFs of presentation did not differ significantly [$F(1,31) = 2.27$, $p = 0.14$, n.s.]

5.5. Event-related brain potential (ERP) results

All p values reported herein are after epsilon correction (Huynh-Feldt) for repeated measures with more than one degree of freedom.

To ensure that our stimulus lateralization manipulation had worked, we analyzed the visual N1 and a selection negativity (100-200 ms and 300-100 ms post noun onset, respectively) at 10 posterior channels. For our primary effects of interest, we analyzed ERPs to target nouns in several time windows, encompassing the frontal N1 (100-200 ms post noun onset at 11 frontal channels), the frontal P2 component (200-300 ms at 11 frontal channels), the N400 component (300-500 ms), and the late positivity (500-1200 ms).

5.5.1. Early ERP effects reflecting stimulus lateralization

5.5.1.1. N1

An early ERP potential, the visual N1, thought to reflect aspects of visual attention (Hillyard & Anllo-Vento, 1998), is known to be larger over the cerebral hemisphere contralateral to the stimulated VF. To ensure that the lateralized presentation of the target nouns effectively stimulated the contralateral hemisphere we examined N1 amplitude (100-200 ms) at the 10 most posterior scalp (non-midline) recording sites. **Figure 5.2** shows ERP responses at a left and a right posterior electrode to targets presented to the two VFs, collapsed across constraint and cloze probability. These measures were subjected to an omnibus 2 (VF: LVF, RVF) X 2 (hemisphere: left, right) x 5 (electrode) ANOVA. The main effects of VF and hemisphere were not significant, but the two factors did interact [$F(1,31) = 72.33$, $p < .0001$], with the expected enhancement in N1 amplitude over left hemisphere scalp sites for RVF presentation , and the opposite pattern for LVF presentation (see **Figure 5.3**).

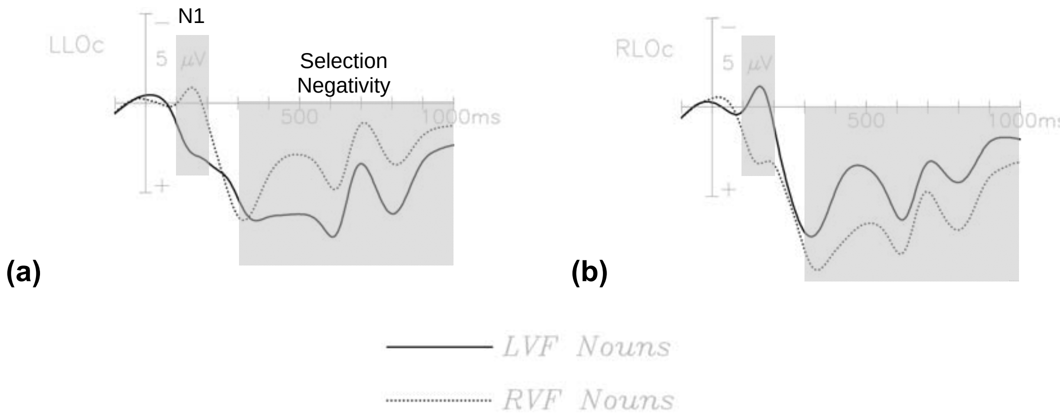


Figure 5.2. Visual N1 (100-200 ms) and selection negativity (300-1000 ms): at (a) a left and (b) a right posterior electrode.

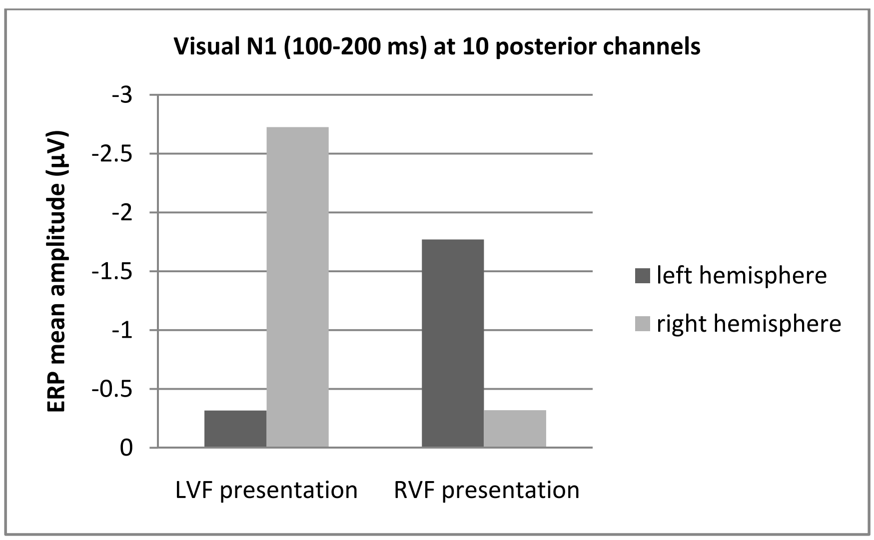


Figure 5.3. Visual N1: Using 100-200 ms time window, 10 posterior electrode sites.

5.5.1.2. Selection negativity

We also noted a contralateral selection negativity – an effect that has similarly been observed in several prior studies using lateralized language stimuli (e.g., Federmeier & Kutas, 1999, 2002; Coulson, Federmeier, Van Petten & Kutas, 2005; Federmeier, Mai & Kutas, 2005; Neville, Kutas & Schmidt, 1982) – in the form of a sustained late negative-going effect over lateral, posterior electrode sites (**Figure 5.2**). To characterize this effect, we took mean amplitude measures between 300-1000 ms from the same 10 posterior electrodes as our visual N1 analysis and subjected them to the same omnibus ANOVA. Like the visual N1, there were no significant effects of either VF or hemisphere, but the two factors interacted [$F(1,31) = 74.59$, $p < .0001$], with LVF stimuli more negative over the right hemisphere sites, and RVF stimuli more negative over left hemisphere sites (**Figure 5.4**).

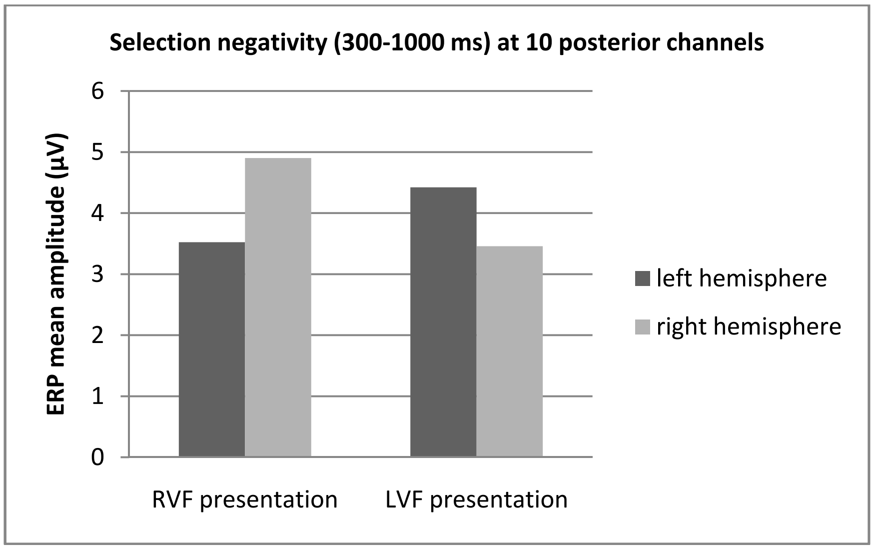


Figure 5.4. Selection negativity: Using 300-1000 ms time window, 10 posterior electrode sites.

5.5.1.3. Summary

Taken together, the modulations in the expected ways of the visual N1 and the selection negativity assured us that our manipulation of stimulus lateralization had been successful.

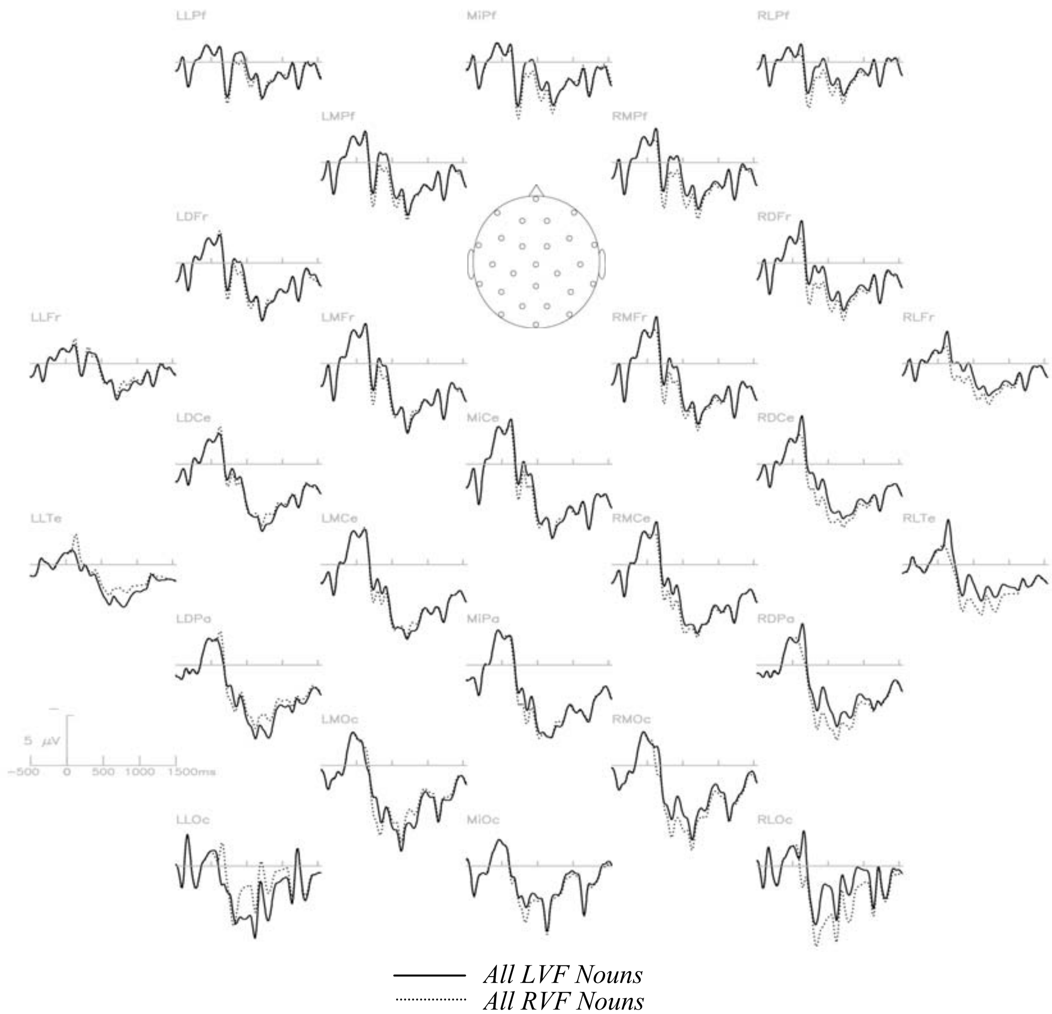
5.5.2. N400: 300-500 ms time window

5.5.2.1. Categorical effects of cloze probability

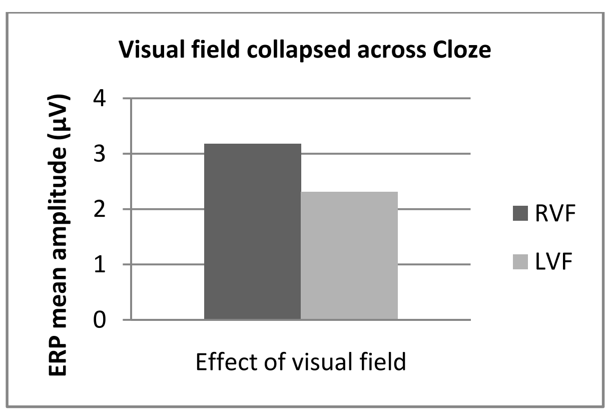
5.5.2.1.1. Overall effect of cloze

To examine whether there was an overall effect of cloze probability in the expected way (larger N400s for lower cloze items), an omnibus ANOVA was conducted with 2 levels of

VF (LVF, RVF) X 2 levels of cloze (<50%, ≥50%) X 26 electrodes. There was a main effect of VF [$F(1,31) = 15.94$, $p = .0004$], with LVF being more negative than RVF (**Figures 5.5a, b**). There was also a main effect of cloze probability [$F(1,31) = 115.15$, $p < .0001$] in the expected direction, with low cloze items being more negative than high cloze items (**Figure 5.6a, b**). There was no significant interaction of VF X cloze (see **Figure 5.7** and **5.8** for cloze effects in individual VFs, and **Figure 5.9** for a comparison of difference waves in the two VFs), nor did VF X cloze X electrode interact. There was an interaction of cloze X electrode, which when explored in further detail with a distributional analysis, revealed a significant interaction of cloze X hemisphere [$F(1,31) = 30.42$, $p = .0000$], with a larger effect of cloze over the right hemisphere (a 2.69µV difference) than left hemisphere (2.20µV difference). Cloze X anteriority also interacted [$F(3,93) = 19.01$, $p = .0001$], with a gradient of cloze effects, increasing in amplitude from anterior to posterior sites. This right, posterior bias of our N400 effect is consistent with the component's "canonical" distributional pattern.

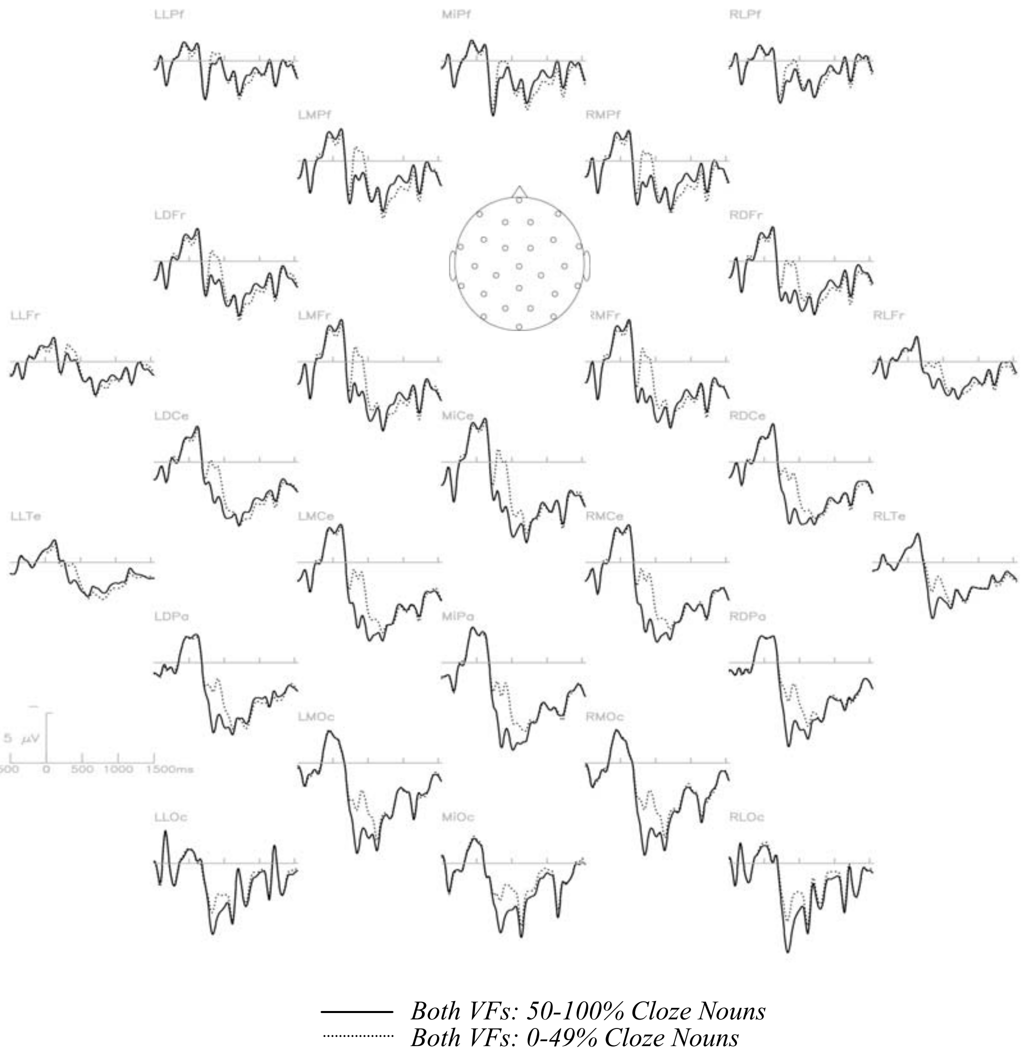


(a)

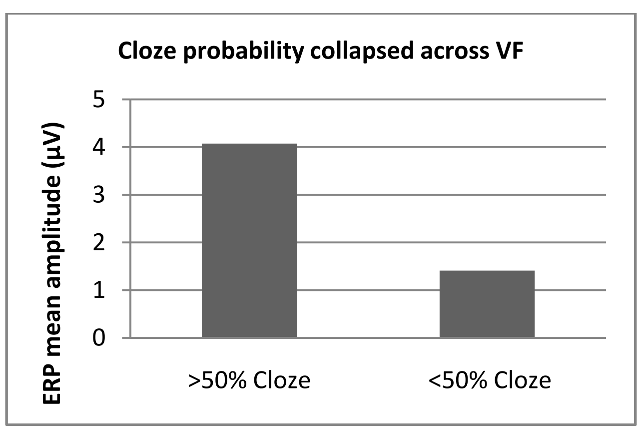


(b)

Figure 5.5. N400 time window (300-500 ms). All LVF vs. All RVF nouns. (a) Plotted over 26 channels and (b) main effect collapsed across electrodes.



(a)



(b)

Figure 5.6. N400 (300-500 ms) time window. Both VFs: high versus low cloze nouns. (a) Plotted over 26 electrode channels and (b) main effect of cloze collapsed across VF.

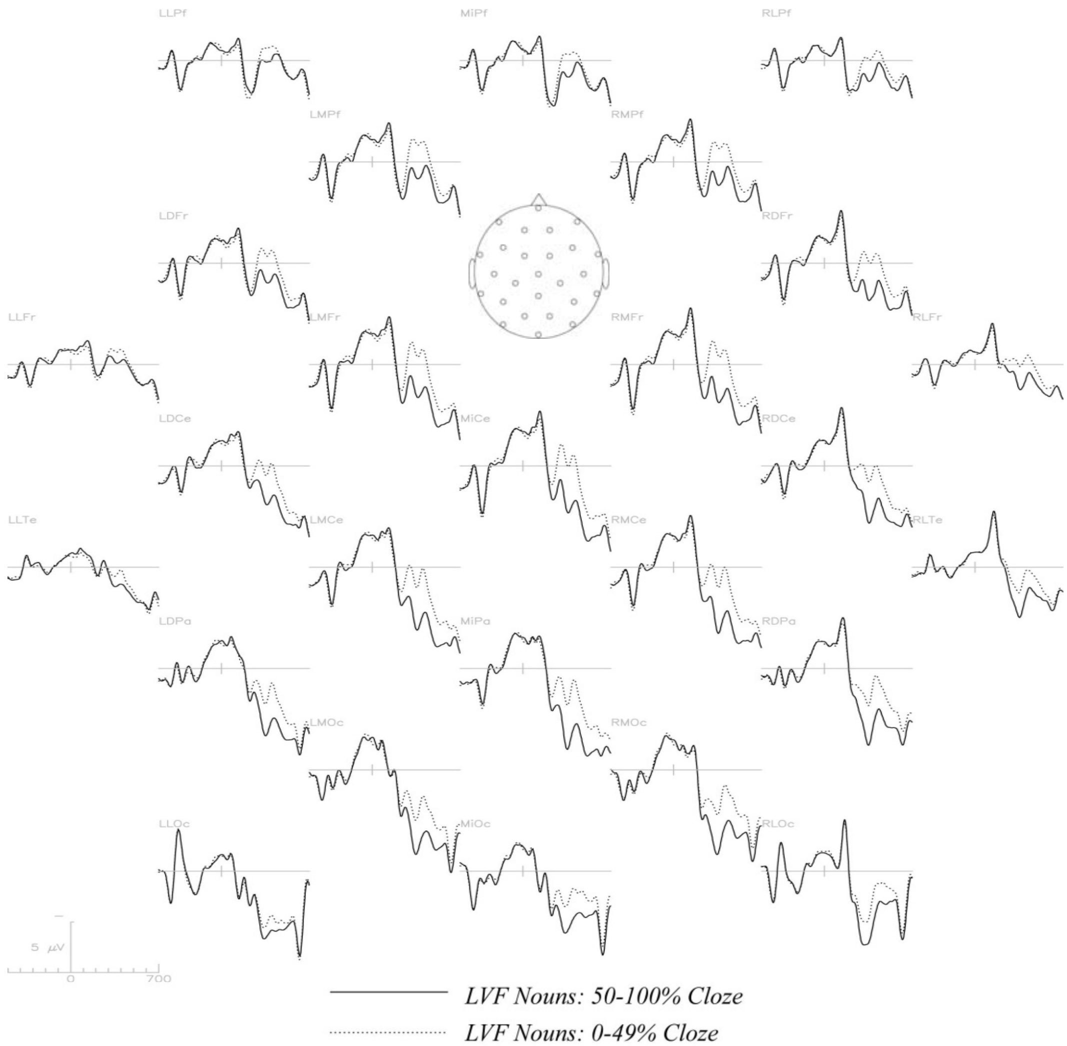


Figure 5.7. LVF: High versus low cloze nouns.

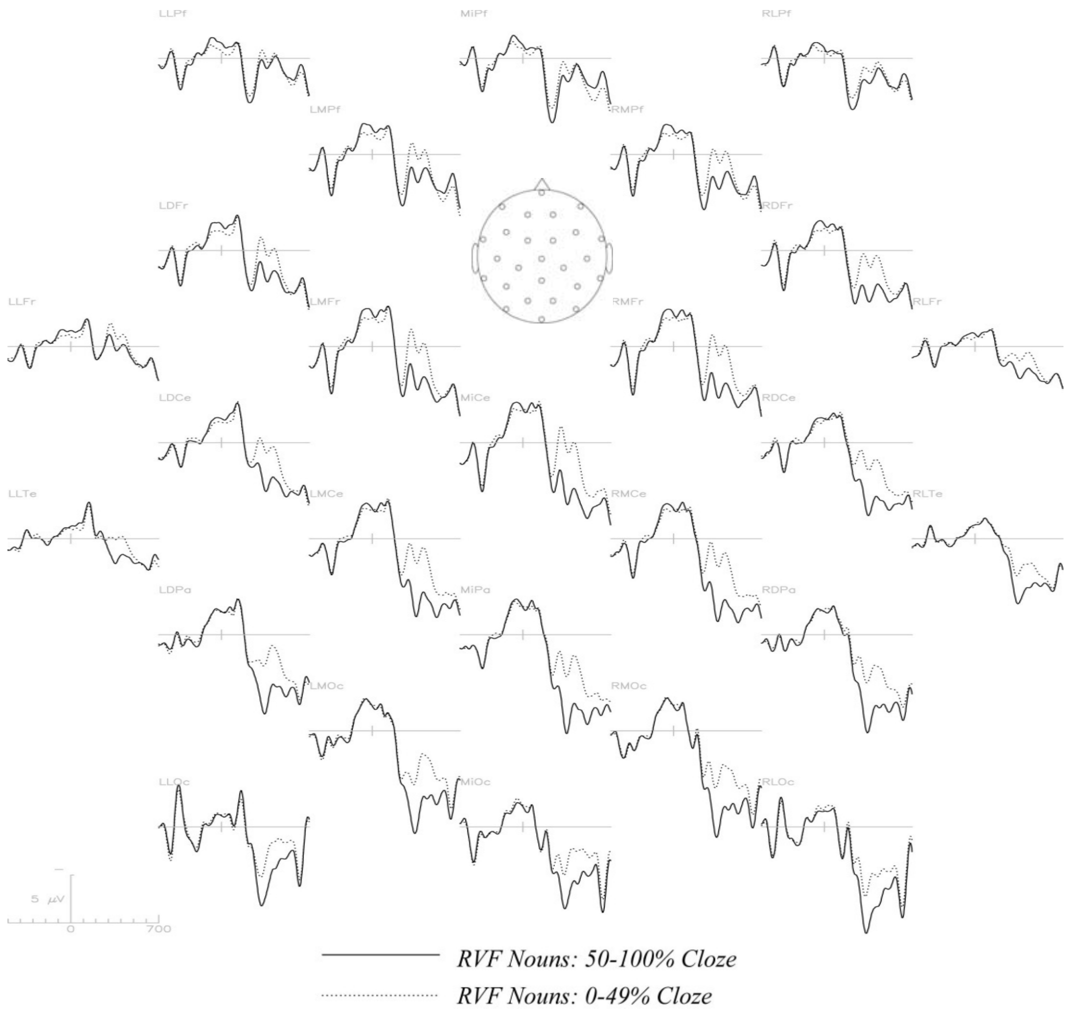


Figure 5.8. RVF: High versus low cloze nouns.

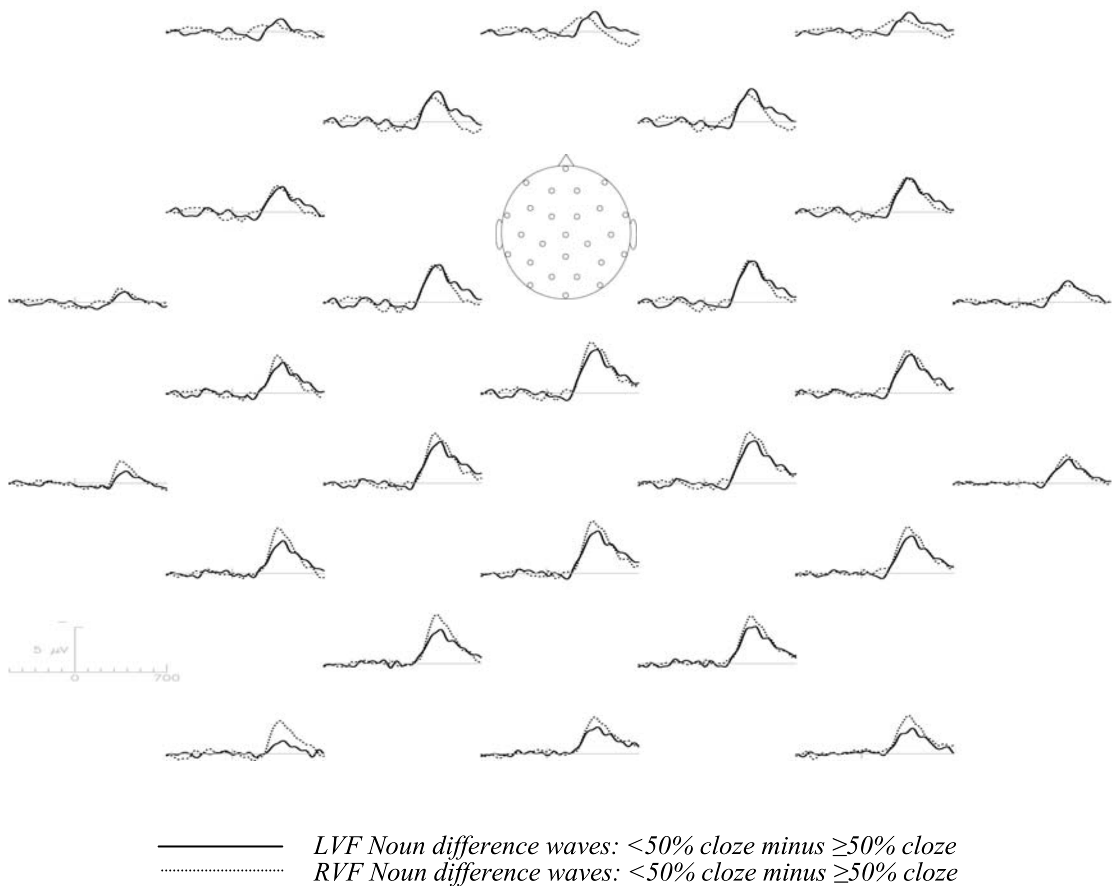


Figure 5.9. Low minus high cloze nouns (difference waves) for comparison of LVF to RVF effects.

To summarize, there was a main effect of cloze (N400 effect) that was not differentiated by VF of presentation. This N400 showed a right posterior bias, a distributional pattern typical of the effect. (There was also a main effect of VF, with LVF being more negative than RVF.)

5.5.2.1.2. Cloze analysis within constraint level

A bar graph showing the mean N400 amplitude for each constraint/cloze condition for each VF of presentation is shown in **Figure 5.10**. This plot is for visualization purposes only, as a statistical analysis including all fourteen constraint/cloze conditions was not conducted due to the difficulty of interpreting potential differences across the combinations of constraint and cloze.

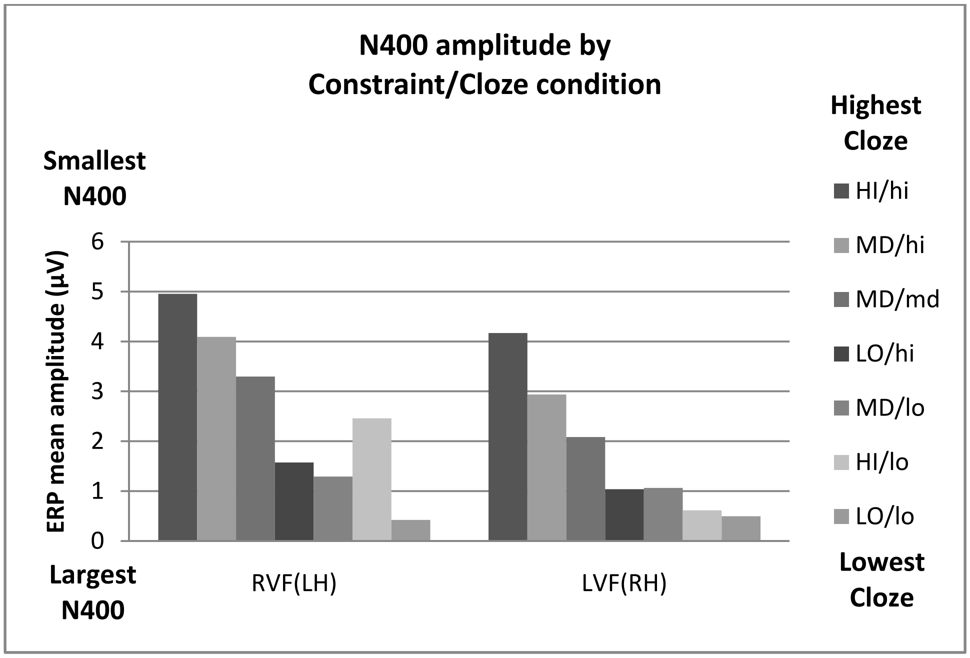


Figure 5.10. N400 time window (300-500 ms). Mean amplitude of seven constraint/cloze conditions by VF.

Instead, these measures were subjected to an omnibus ANOVA at each level of constraint (HI, MD, and LO): 2 levels of VF (LVF vs. RVF) X 2[3, for MD constraint] levels of cloze probability (hi [vs. md] vs. lo cloze) X 26 electrodes.

Within HI CONSTRAINT: VF(2) X CLOZE(2) X ELECTRODE(26): For the HI constraint condition: (a) there was a main effect of VF [$F(1,31) = 12.45$, $p = .0013$], with responses to LVF words being more negative than those for the RVF (1.31 μV difference), (b) there was a main effect of cloze probability [$F(1,31) = 44.27$, $p < .0001$], with larger N400 amplitudes to lo cloze (1.54 μV) compared to hi cloze (4.56 μV) words, and (c) visual field and cloze did not interact, either over all 26 electrode sites or distributionally. See **Figure 5.11a** and **b** for HI constraint (high vs. low cloze) ERPs within each VF, and as difference waves (**Figure 5.11c**).

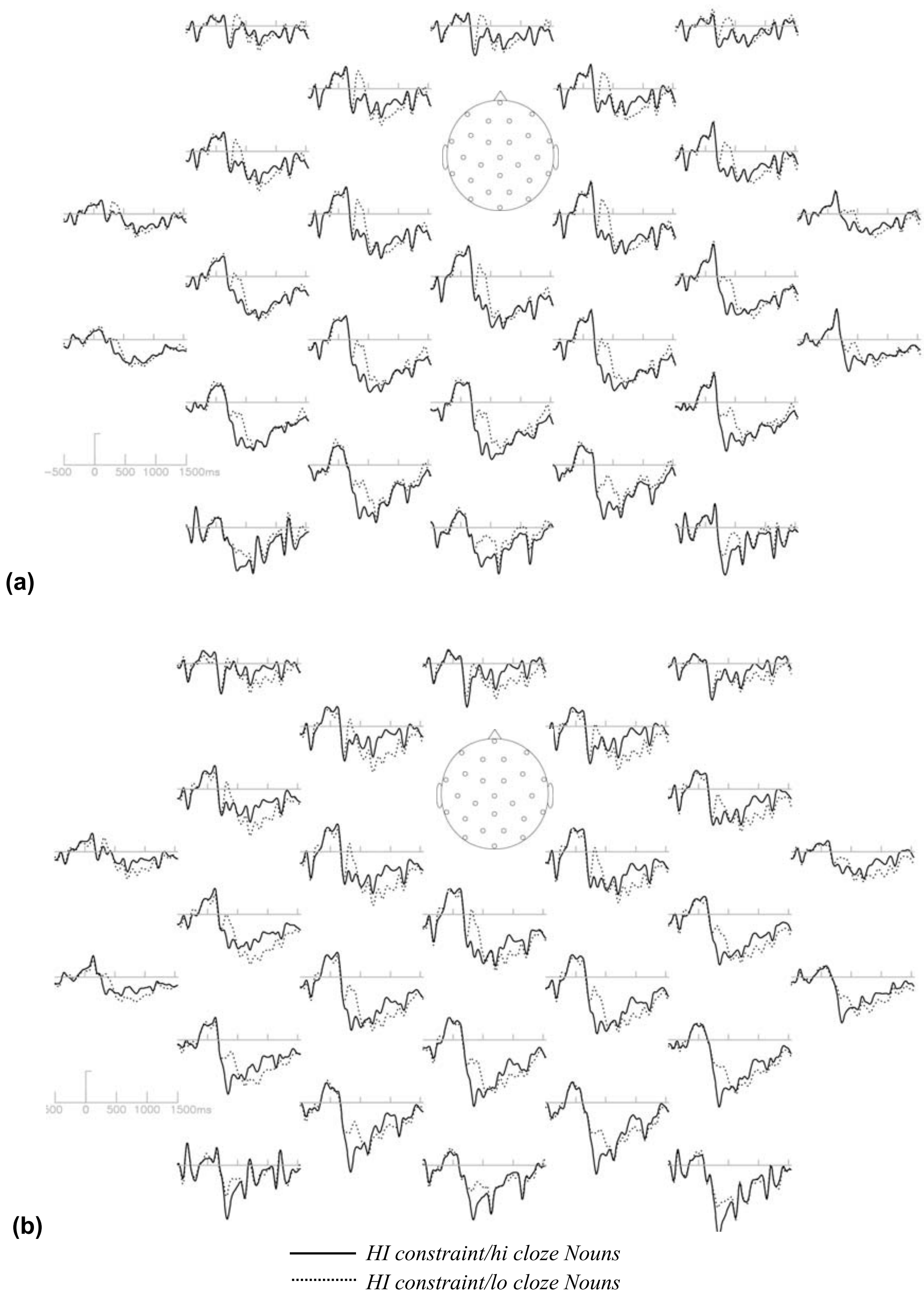
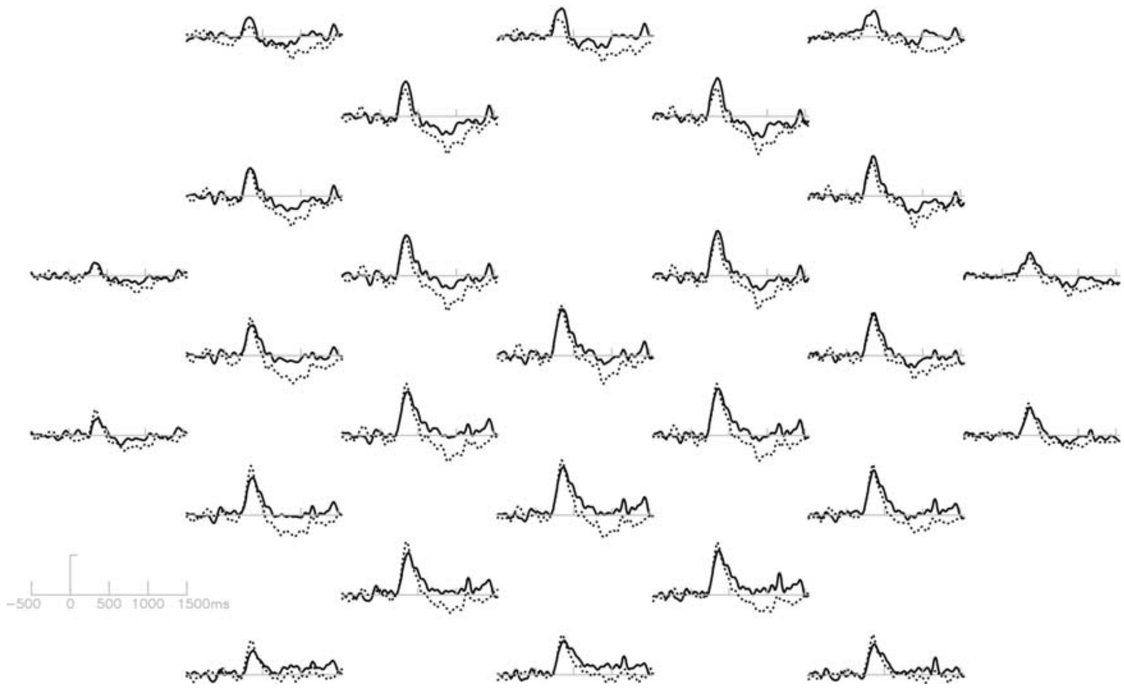


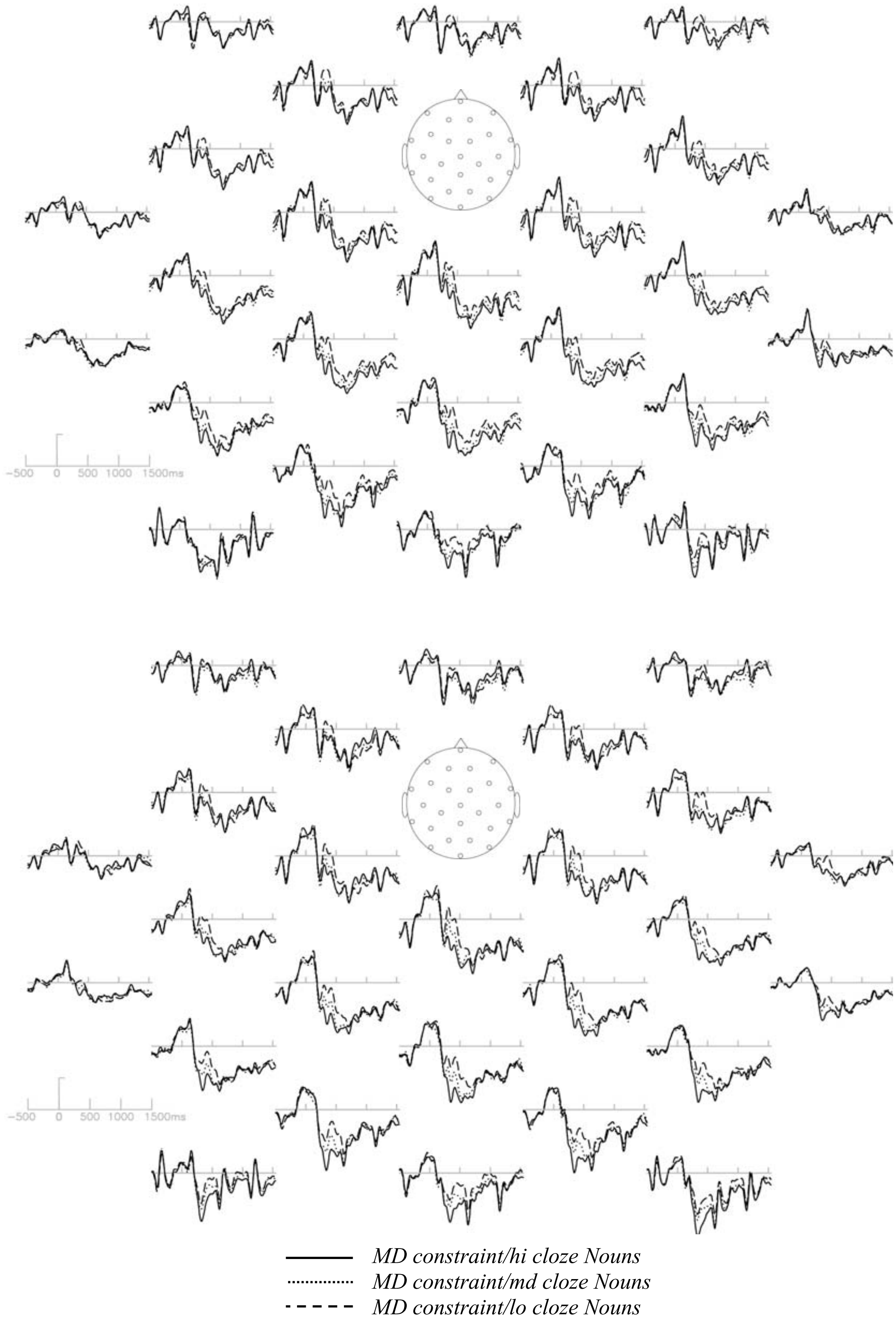
Figure 5.11. HI constraint nouns: high vs. low cloze for (a) LVF presentation and (b) RVF presentation.



LVF: HI constraint/low minus high cloze Noun difference wave
RVF: HI constraint/low minus high cloze Noun difference wave

Figure 5.11. (c) HI constraint nouns: low minus high cloze (difference waves) for comparison of LVF to RVF effects.

Within MD CONSTRAINT: VF(2) X CLOZE (3) X ELECTRODE(26): For the MD constraint condition: (a) there was a main effect of visual field [$F(1,31) = 7.91$, $p = .0084$], with responses to LVF words being more negative than those for the RVF (.87 μV difference), (b) there was a main effect of cloze probability [$F(2,62) = 50.72$, $p < .0001$], with larger N400 amplitudes to lo cloze (1.18 μV) compared to md (2.69 μV) compared to hi cloze (3.52 μV) words, and (c) visual field and cloze did not interact, either over all 26 electrode sites or distributionally. See **Figure 5.12a** and **b** for MD constraint (hi vs. md vs. lo cloze) ERPs within each VF.

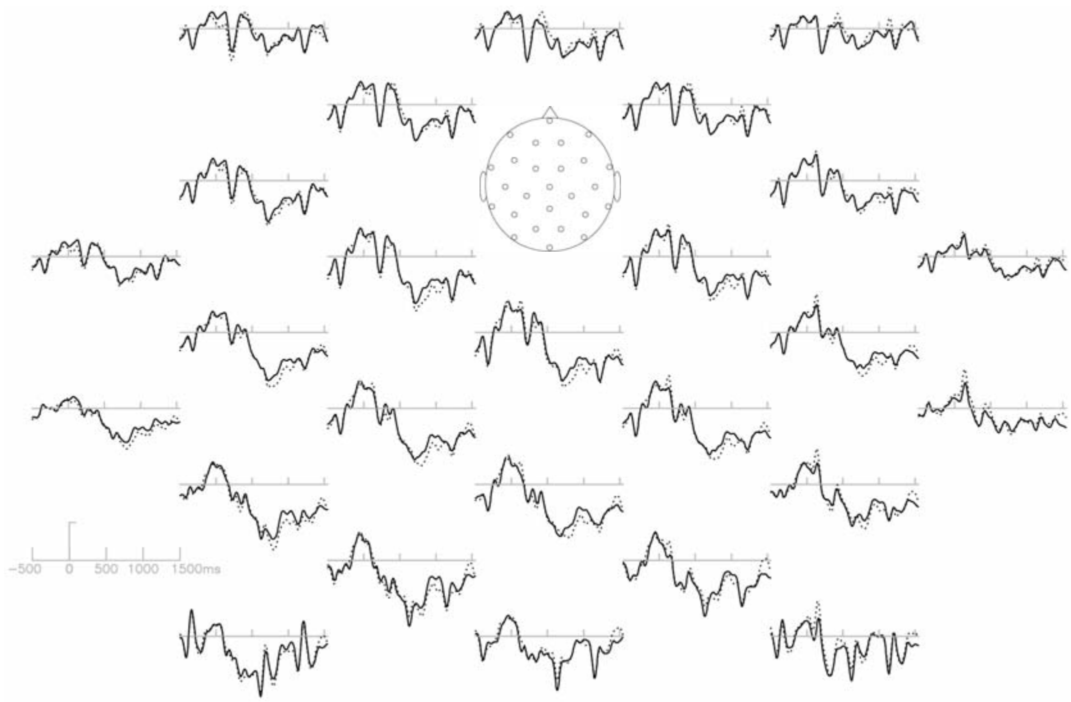


(a)

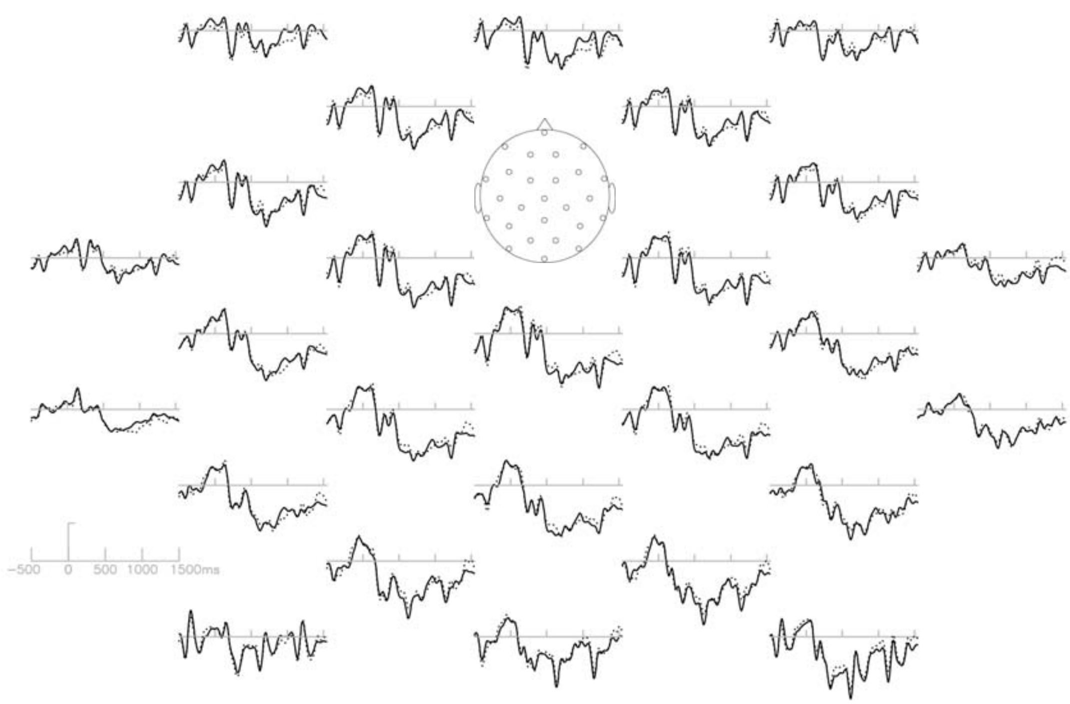
(b)

Figure 5.12. MD constraint nouns: high vs. medium vs. low cloze for (a) LVF presentation and (b) RVF presentation.

Within LO CONSTRAINT: VF(2) X CLOZE (2) X ELECTRODE(26): For the LO constraint condition: (a) there was no main effect of visual field [$F(1,31) = .40$, $p = .5320$], (b) there was a main effect of cloze probability [$F(1,31) = 6.23$, $p = .0181$], with larger N400 amplitudes to low cloze (.46 μV) compared to high cloze (1.31 μV) words, and (c) visual field and cloze did not interact, either over all 26 electrode sites or distributionally. See **Figure 5.13a** and **b** for LO constraint (high vs. low cloze) ERPs within each VF.



(a)



(b)

LO constraint/hi cloze Nouns
LO constraint/lo cloze Nouns

Figure 5.13. LO constraint nouns: high vs. low cloze for (a) LVF presentation and (b) RVF presentation.

To summarize (**Table 5.5**), there were main effects of cloze within each of the three constraint levels, in the expected direction, with similar effects over the two visual fields. (For HI and MD constraint contexts, there was also a main effect of VF, with LVF being more negative than RVF.)

Table 5.5. Summary of Noun N400 cloze analyses within constraint levels.

Constraint	**HI**	**MD**	**LO**
Main effect of Visual Field	Yes, LVF more negative than RVF	Yes, LVF more negative than RVF	No
Main effect of cloze probability across both Visual Fields	Yes, typical N400	Yes, typical N400	Yes, typical N400
Interaction of Visual Field X cloze probability	No	No	No
Interaction of Visual Field X cloze probability X electrode	No	No	No

5.5.2.1.3. Cloze analysis within each VF, within each constraint level

Also, within each visual field, planned comparisons were made for cloze probability (hi/lo for HI and LO constraint, and hi/md/lo for MD constraint) within each constraint level. The results are shown in **Table 5.6**, with interactions of cloze X electrode followed up with distributional analyses.

Table 5.6. Cloze probability findings within constraint levels, Noun N400 time window (300-500 ms). Significant findings are highlighted in bold.

		RVF (LH)	**LVF (RH)**
HI Constraint (2 levels: hi, lo)	Main Effect of cloze	***F*(1,31) = 16.86, *p* =.0003***	***F*(1,31) = 45.90, *p* =.0000***
	Interaction of cloze with electrode	***F*(25,775) = 9.60, *pHF* =.0000***	***F*(25,775) = 8.64, *pHF* =.0000***
MD Constraint (3 levels: hi, md, lo)	Main Effect of cloze	***F*(2,62) = 19.62, *pHF* =.0000***	***F*(2,62) = 17.42, *pHF* =.0000***
	Interaction of cloze with electrode	***F*(50,1550) = 4.67, *pHF* =.0002***	***F*(50,1550) = 4.67, *pHF* =.0000***
LO Constraint (2 levels: hi, lo)	Main Effect of cloze	***F*(1,31) = 6.26, *p* =.0178***	*F*(1,31) = 1.17, *p* =.2872, n.s.
	Interaction of cloze with electrode	*F*(25,775) = 1.96, *pHF* =.1045, n.s.	*F*(25,775) = .39, *pHF* =.7616, n.s.

5.5.2.1.3.1. Distributional effects for LVF(RH) presentation

Within HI constraint contexts: There was an interaction of cloze X hemisphere [$F(1,31) = 7.05$, $p = .0124$], with a larger effect of cloze over the right hemisphere (a 3.56μV difference) than left hemisphere (3.00μV difference). There was also an interaction of cloze X laterality [$F(1,31) = 23.77$, $p < .0001$] with larger cloze effects over medial sites (4.29μV) than lateral sites (2.27μV). These interactions were mediated by a higher order interaction of cloze X hemisphere X laterality [$F(1,31) = 4.40$, $p = .0441$], with left and right medial sites exhibiting more similar cloze effects (only a 0.32μV difference) than at lateral sites, where left lateral effects were smaller than right lateral effects by 0.80μV. Cloze X laterality X anteriority also interacted significantly [$F(3,93) = 6.49$, $pHF = .0026$], with the largest cloze effects (a 5.05μV effect) over centro-parietal sites at medial scalp locations but over occipital sites (a 3.10μV effect) at the lateral scalp locations.

Within MD constraint contexts: There was an interaction of cloze X hemisphere [$F(2,62) = 3.80$, $pHF = .0276$], with larger cloze effects over the right than over the left hemisphere. There was also an interaction of cloze X laterality [$F(2,62) = 9.01$, $pHF = .0004$], with larger cloze effects over medial than lateral sites. These interactions were mediated by a higher order interaction of cloze X hemisphere X laterality [$F(2,62) = 5.51$, $pHF = 0.0063$] which revealed that the cloze effect at left lateral sites was reduced (a .57uV difference between high and low cloze) compared to that at right lateral sites (1.52uV effect), and relative to both left and right medial sites, (2.22uV and 2.53uV effects, respectively). There was also an interaction of cloze X anteriority [$F(6,186) = 4.48$, $pHF = 0.0058$] indicating that cloze effects progressively decreased in amplitude from posterior to anterior scalp sites.

5.5.2.1.3.2. Distributional effects for RVF(LH) presentation

Within HI constraint contexts: There was an interaction of cloze X laterality [$F(1,31) = 13.86$, $pHF = 0.0008$] revealing larger cloze effects at medial than lateral sites, and an interaction of cloze X anteriority [$F(3,93) = 11.15$, $pHF = 0.0019$] revealing progressively smaller cloze effects from posterior to anterior sites.

Within MD constraint contexts: There were interactions of cloze X hemisphere and cloze X laterality, which were mediated by a higher order interaction of cloze X hemisphere X laterality [$F(2,62) = 4.40$, $pHF = 0.0164$]. This interaction indicated that the cloze effect at left lateral sites was reduced (a .1.27µV difference between high and low cloze) compared to that at right lateral sites (2.11uV effect), as well as at left and right medial sites (3.33µV and 3.60µV effects, respectively). There was also a significant interaction of cloze X laterality X anteriority [$F(6,186)= 3.28$, $pHF = 0.0240$] suggesting that differences in the amplitude of cloze effects between lateral and medial sites were greatest at frontal and centro-parietal sites, compared to those at pre-frontal and occipital sites.

Within LO constraint contexts: There was a significant interaction of cloze X laterality which was mediated by a higher order interaction of cloze X hemisphere X laterality [$F(1,31) = 6.80$, $pHF = 0.0139$]. These interactions indicated that while in general the N400 effect was larger over medial than lateral sites, this was primarily due to the contribution of cloze effects at right, but not left, lateral sites.

5.5.2.1.3.3. Summary

N400 effects were significant over both VFs, within all constraint levels except LVF/Lo constraint contexts, where there was no significant difference between LO/hi and LO/lo contexts. Additionally, each of the significant cloze effects within VF and constraint level interacted with electrode site. Overall, these distributional effects indicated that N400 effects were larger over the right than left hemisphere, larger over medial than lateral sites, and were also more posterior than anterior. Interactions also revealed that cloze effect size generally diminished over left lateral sites.

5.5.2.2. Correlations of cloze probability and N400 amplitude

In addition to traditional ANOVAs, correlations of target noun N400 mean amplitude and cloze probability were conducted, collapsing across level of constraint. The individual nouns were sorted into ten equal-width bins as a function of cloze probability, from highest (90–100%) to lowest (0–10%). ERPs for each 10% bin were averaged first within, then across, participants. The average numerical cloze probability of each bin was then calculated and correlated with mean ERP amplitude in the N400 time window (300–500 ms). Correlation coefficients (*r*-values) were then calculated separately for all 26 electrode sites. These values are plotted separately for VF of presentation in **Figure 5.14**.

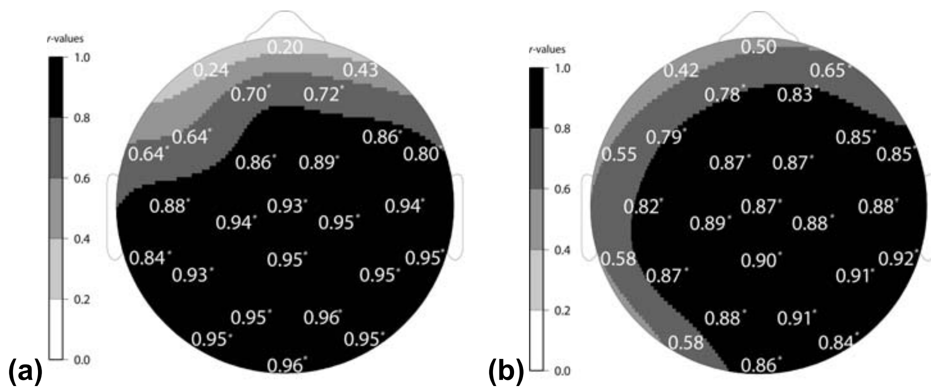


Figure 5.14. Noun cloze X Mean Amplitude Correlations (*r*-values), N400 time window (300-500 ms). (a) RVF(LH) and (b) LVF(RH). Significant correlations ($p \leq .05$) are indicated with asterisks (*).

To determine whether N400 mean amplitude was linearly correlated with cloze probability more for one visual hemifield than the other, the following analyses were conducted: because the data were highly correlated (with potentials across most of the electrodes behaving similarly), principal components analysis was used to reduce the dimensionality by first removing the mean of the data and then compressing the data to the first principal component (PC). The first PC alone accounted for 90-93% of the data variance. A standard one-dimensional regression run on the subsequent data revealed that for presentation to both hemifields, N400 amplitude highly co-varied with cloze, $r = 0.94$, $p < .00001$ for RVF(LH) and $r = 0.91$, $p < .001$ for LVF(RH). To test that the linear correlation coefficients from the two samples were equal, a Fisher *z*-score transformation of the correlation coefficients was performed, revealing that neither hemisphere exhibited a stronger relationship with cloze than the other ($p = 0.65$, n.s.).1

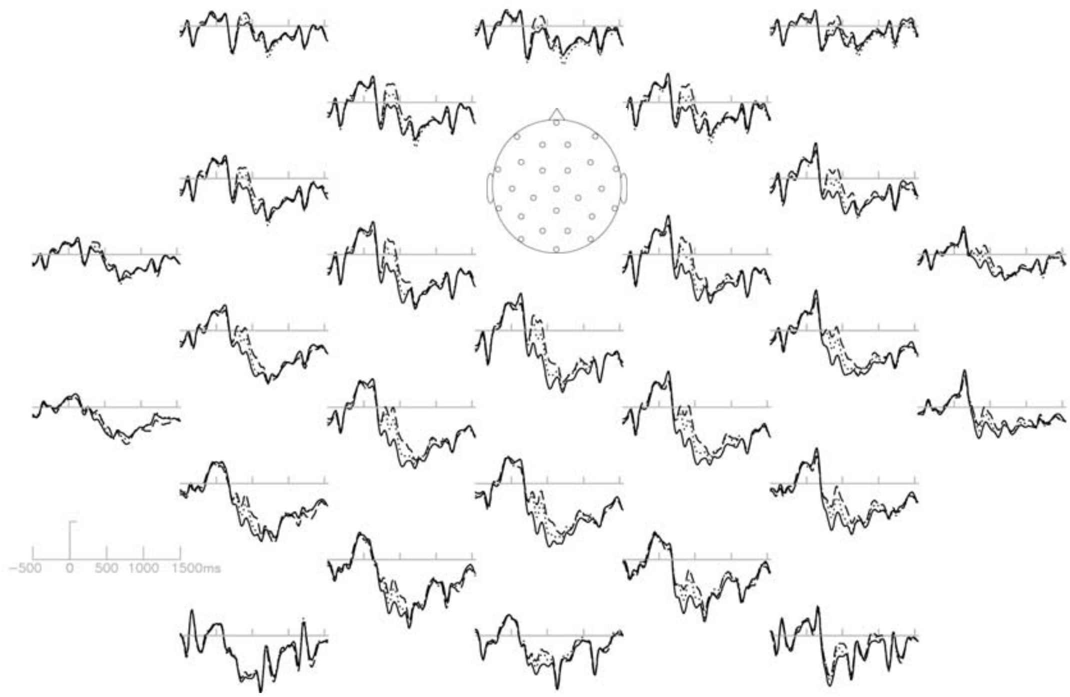
In sum, the relationship of cloze with N400 mean amplitude was similar for both hemifields, with widespread, comparably high *r*-values (.80s-.90s), and right, posterior maximal correlations for both VFs.

5.5.2.3. Effects of constraint

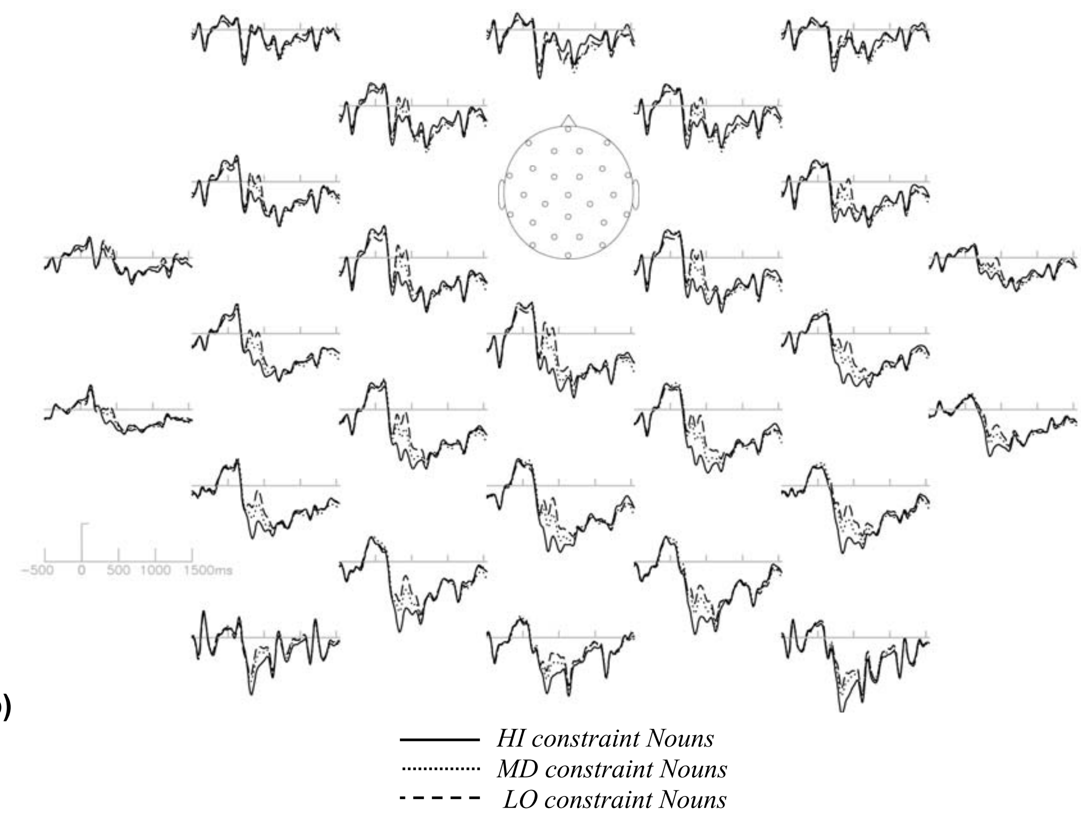
5.5.2.3.1. Overall effect of constraint

For constraint, an omnibus ANOVA with 2 levels VF (LVF, RVF) X 3 levels constraint (HI, MD, LO) X 26 levels electrode was conducted. Based on the results from the cloze analysis and the knowledge that factors of constraint and cloze are confounded, an overall main effect of constraint was expected, though would not be very explanatory. In addition to a significant main effect of VF, with the LVF being more negative than RVF by .81 μV, [$F(1,31) = 12.26$, $p = .0014$], there was indeed a significant main effect of constraint in the expected direction, [$F(1,31) = 55.16$, $p < .0001$], with HI (3.76 μV) > MD (2.21 μV) > LO (1.08 μV) constraint. There was not, however, a significant interaction between VF X constraint [$F(2,62) = 1.46$, $p = .2401$, n.s.]. See **Figure 5.15a** and **b** for HI vs. MD. vs. LO constraint nouns.

[1] The principal component analyses (PCA) performed to test the similarity of the LVF and RVF N400 amplitude/cloze relationship was contributed by David M. Groppe.



(a)



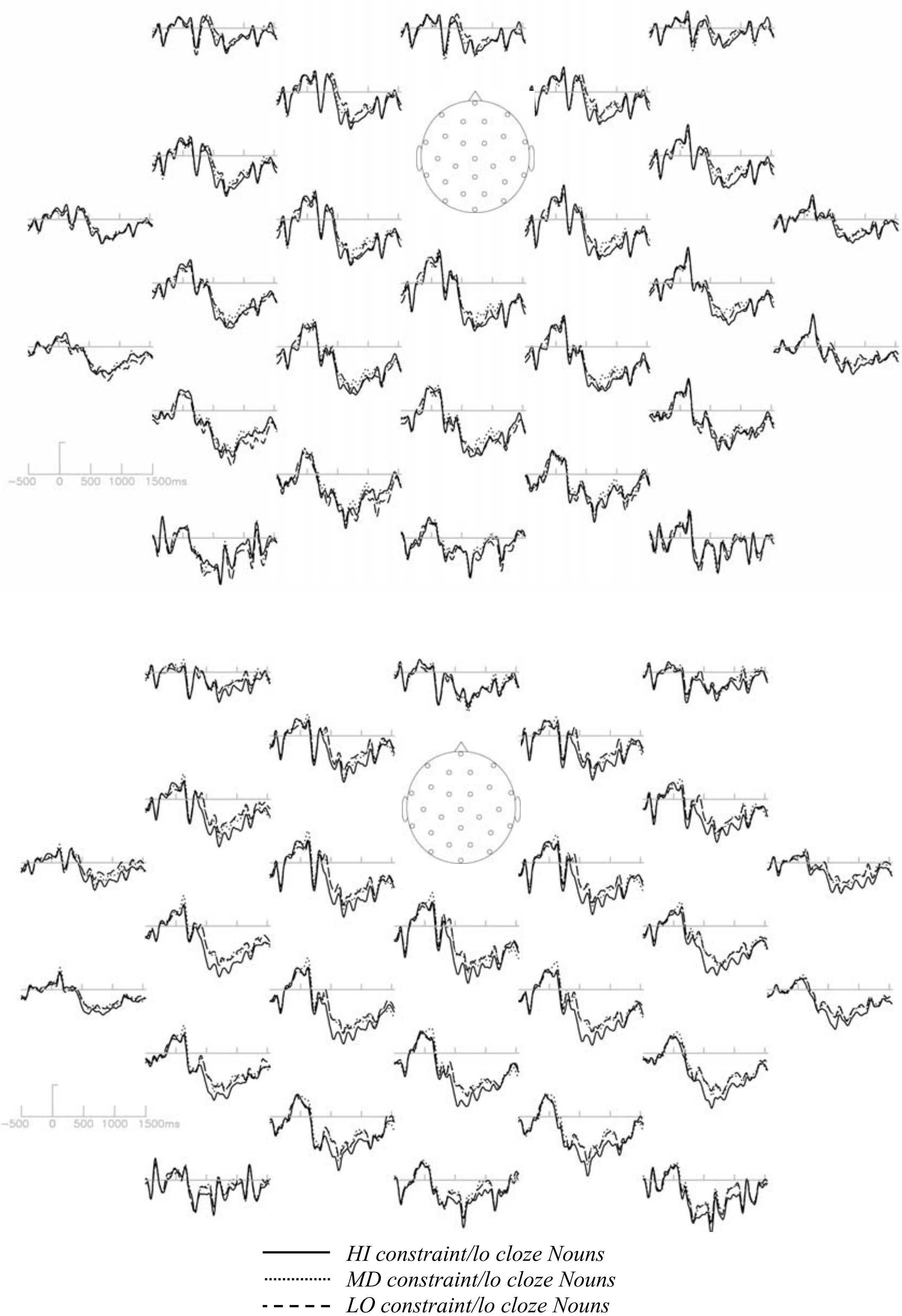
(b)

Figure 5.15. HI vs. MD vs. LO constraint nouns. (a) LVF presentation and (b) RVF presentation.

In sum, though this analysis is not particularly informative (due to the confound of constraint with cloze measures), there was an expected main effect of constraint, with LO constraint continuations more negative than MD constraint, which were more negative than HI constraint nouns. (There was also a main effect of VF, with LVF being more negative than RVF.)

5.5.2.3.2. Effects of constraint violation and visual field: Using only low cloze conditions (HI/lo, MD/lo, LO/lo)

A more informative way to examine effects of constraint violation is to contrast only the low cloze conditions within each level of constraint (HI,MD,LO), because here constraint level can be manipulated while cloze probability is held relatively constant (with mean values of cloze probability for the three low cloze conditions being .04, .05, and .02 respectively). In the following analysis, then, we compare how constraint violation was impacted by visual field, contrasting only the low cloze conditions of each level of constraint (HI,MD,LO). **Figure 5.16a** and **b**.



(a)

(b)

Figure 5.16. HI,MD,LO constraint/low cloze nouns. (a) LVF presentation and (b) RVF presentation.

Using the following three constraint/cloze conditions (HI/lo, MD/lo, LO/lo) X 2 levels VF (LVF, RVF) X 26 levels electrode, within the N400 time window there was a main effect of visual field [$F(1,31) = 5.19$, $p = .03$], with responses to LVF words overall being more negative than those for the RVF (.6665 μV difference). There was also a main effect of constraint violation [$F(2,62) = 5.21$, $pHF = .0084$], with HI/lo, MD/lo, and LO/lo constraints exhibiting the following mean amplitudes, respectively: 1.52μV, 1.18 μV, and 0.46μV (see **Figure 5.17** below). However, this effect of constraint violation was mediated by an interaction with visual field [$F(2,62) = 4.08$, $pHF = .0222$] (see **Figure 5.17** below), analyzed more fully by conducting a planned comparison of the three constraint/cloze conditions within each visual field.

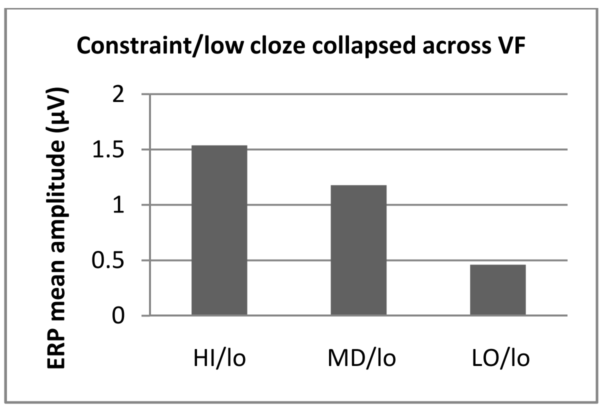


Figure 5.17. N400 time window. Constraint/lo cloze collapsed across VF.

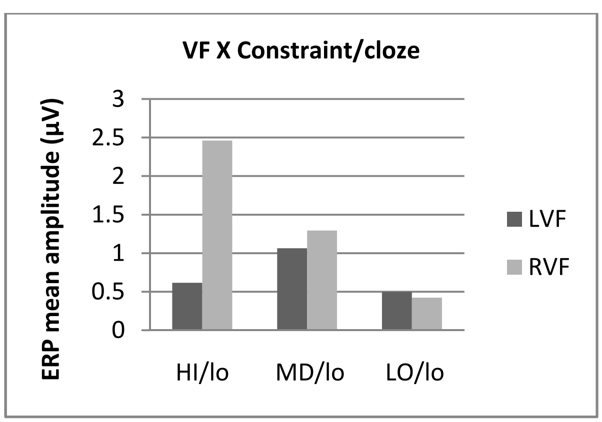


Figure 5.18. N400 (300-500 ms) time window. VF X constraint/lo cloze.

An omnibus ANOVA (3 levels of constraint/cloze X 26 levels of electrode) for each visual field revealed a significant main effect of constraint for the RVF [$F(2,62) = 9.83$, $pHF = .0002$], but not LVF [$F(2,62) = 0.64$, $pHF = .5309$, n.s.], with the following respective HI/lo, MD/lo, and LO/lo condition mean amplitude values: 2.46, 1.29, 0.42 µV. For the RVF there was also an interaction of constraint/cloze and electrode site: [$F(50,1550) = 2.20$, $pHF = .0357$]. Distributional analyses revealed a significant interaction of constraint/cloze X hemisphere X laterality [$F(2,62) = 3.63$, $pHF = .03$], with left/lateral sites showing the smallest amplitude constraint violation effects:

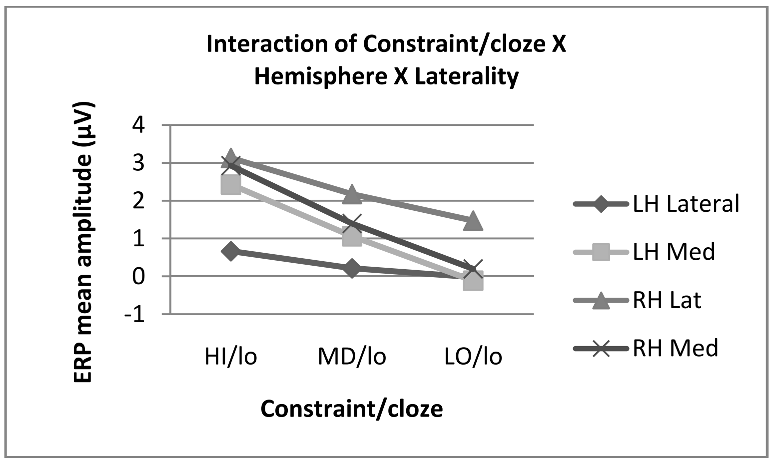


Figure 5.19. N400 (300-500 ms) time window. For RVF(LH), constraint/lo cloze X hemisphere X laterality.

5.5.2.3.2.1. Constraint violation effects for RVF(LH) only: 3 levels constraint/cloze (HI/lo, MD/lo, LO/lo) X 2 levels time window (300-400 ms, 400-500 ms) X 26 levels of electrode

As the presence of a constraint violation effect during the N400 time window differed from the results previously reported for centralized presentation (Kutas & Hillyard, 1984), we wanted to further investigate the increased positivity between 300-500 ms for RVF presentation. If these results reflect something systematic about N400 effects, per se, in relation to constraint violation, we might expect to observe consistently large mean amplitude patterns throughout the entire N400 time window. However, another possibility is that the RVF results are reflecting an earlier-than-expected onset of the LP, and if this is the case, we might expect larger effects of constraint violation in the latter, relative to the earlier, portion of the N400 time window. To investigate these possibilities we divided the N400 time window into an earlier (300-400 ms) and later (400-500 ms) portion and conducted an ANOVA using the three levels of constraint violation (HI/lo, MD/lo, LO/lo) and two levels of time window. While results indicated an overall effect of increasing ERP positivity with constraint violation [$F(2,62) = 9.83$, $pHF = .0002$], an interaction of constraint violation with time window [$F(2,62) = 5.20$, $pHF = .0082$] indicated a significantly larger effect over the latter half of the N400 time window. See **Figure 5.20**. These results suggest that the effect of constraint violation within the N400 time window reflects the overlap of the LP component within the N400 time window, rather than indicating that the constraint violation is modulating the N400 component directly.

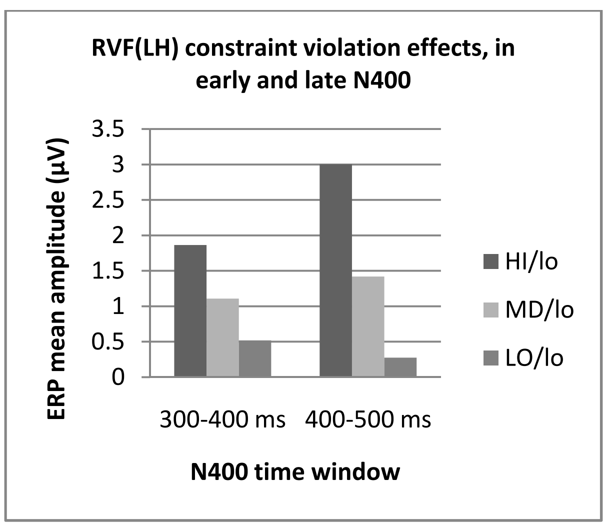


Figure 5.20. Early and late N400 time window. RVF(LH) effects of constraint violation.

5.5.2.3.2.2. Summary of constraint violation effects by VF

In sum, in addition to there being a main effect of VF (with LVF being more negative than RVF), over both VFs there was a main effect of constraint, with low cloze continuations for HI constraint contexts more positive than those for MD constraint, which were more positive than those for LO constraint. However, an interaction of constraint and VF revealed that this relationship held only for the RVF, not the LVF. This effect was present over all but the left lateral sites. This finding differs from previous findings, where constraint violation did not modulate ERP mean amplitude in the N400 time window. However, it also runs contrary to what would have been considered the expected direction of such a finding: instead of increased negativity in the N400 time window as constraint violation increased, there was increased *positivity* with increased constraint violation. This effect of constraint violation for RVF(LH) appears to reflect the overlap of an (earlier-than-expected) LP within the N400 time window.

5.5.2.4. Correlations of constraint violation and N400 mean amplitude

Based on the above ANOVAs, correlations were also conducted for the low cloze noun continuations of all three constraint conditions (HI/lo, MD/lo and LO/lo) so that in a more fine-grained manner, it could be determined whether the mean amplitude pattern in the N400 time window varied systematically with degree of contextual constraint violation. Based on previous studies (e.g., Kutas & Hillyard, 1984) constraint did not modulate ERP mean amplitude to low cloze sentence endings in the N400 time window. But since our ANOVA analysis did not pattern in a similar way to the centrally presented Kutas & Hillyard study (with our results showing that within the N400 time window, low cloze continuations to high constraint contexts were more positive than those to low cloze continuations to low constraint contexts, at least for RVF presentation), we were motivated to further explore the results through correlation analysis. There are three possible results for a constraint violation correlational analysis in the N400 time window (**Figure 5.21**): (a) constraint violation and N400 amplitude may not be significantly correlated, (b) a pattern of positive correlations would indicate that the more constraint is violated, the greater the positivity, or (c) a pattern of negative correlations would indicate that the more constraint is violated, the greater the negativity (N400) – a finding which would contradict previous evidence for constraint violation not modulating N400 amplitude.

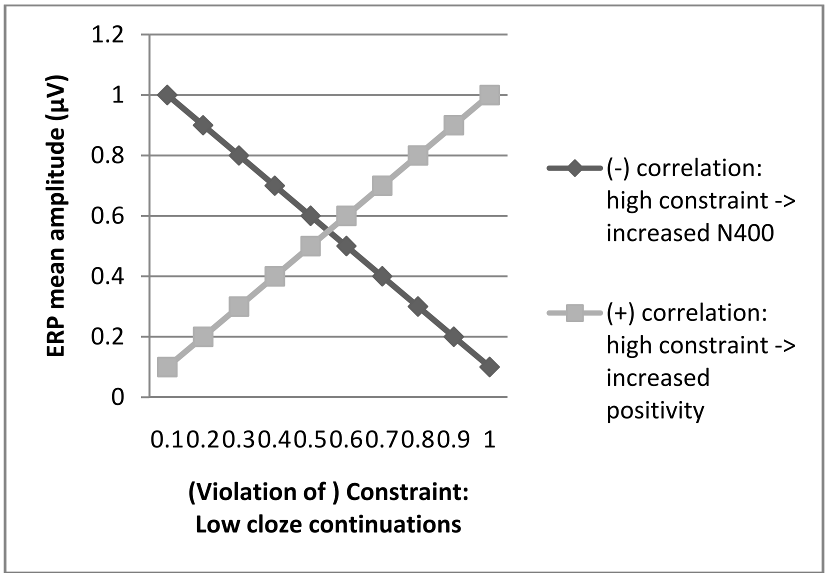


Figure 5.21. Possible correlation mappings of violation of constraint correlated with mean amplitude in N400 time window (300-500 ms).

To conduct this analysis, HI/lo, MD/lo, and LO/lo targets were sorted by their contextual constraint values into 5 20% bins (**Table 5.7**), and ERP mean amplitude and constraint violation were correlated within the N400 (300-500ms) time window.

Table 5.7. HI, MD, LO constraint/low cloze continuations sorted by contextual constraint.

Constraint Percent Bin	Number of Items	Mean Bin Constraint
82-100%	35	92%
60-81%	7	79%
40-59%	22	47%
20-39%	65	28%
0-19%	48	13%

The correlation patterns of mean amplitude of low cloze items correlated with constraint level in the N400 time window indicated a pattern of strong positive correlations (as constraint violation increases, ERP mean amplitude becomes more positive), particularly for RVF(LH) presentation (see **Figure 5.22**).

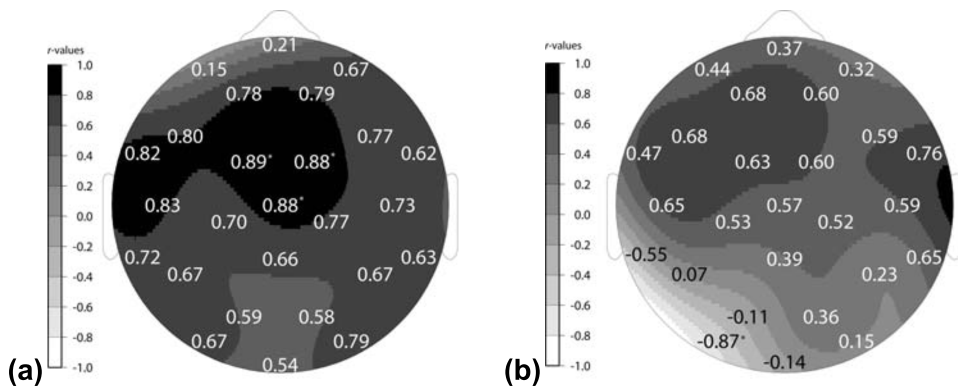


Figure 5.22. Correlations of Mean Amplitude in N400 Time Window (300-500ms) with Low cloze items sorted on constraint of Context (HI+MD+LO/lo items), (a) RVF(LH) and (b) LVF(RH). Significant correlations (p≤.05) are indicated with asterisks (*).

In sum, this analysis revealed that as constraint violation increased, ERP mean amplitude for lo cloze nouns became more positive. This pattern was more evident (stronger correlations) for RVF than LVF presentation, with maximal *r*-values primarily over the left, anterior scalp quadrant.

5.5.3. Late positivity (LP): 500-1200 ms time window

Based on visual observation of the waveforms in the present study, as well as previous studies conducted using the same stimuli, a rather extensive time window, 500-1200 ms, was used to examine the late positivity (LP). Though shorter time windows are often used to characterize P600 effects and other LPs (e.g., 500-800 or 500-900 ms), across studies the duration of LP time window tends to be much more variable than, for instance, that of the N400.

5.5.3.1. Effects of cloze probability

5.5.3.1.1. Overall effect of cloze

To examine whether there was an overall effect of cloze probability in the late time window, an omnibus ANOVA was conducted with 2 levels of VF X 2 levels of cloze X 26 electrodes. Of interest here were potential effects of cloze and interactions between VF X cloze, so distributional effects of VF alone will not be reported. There was no main effect of cloze (refer to **Figure 5.6a** for ERP plots), nor were there any significant interactions of cloze X VF (**Figures 5.7**and **5.8**). There was, however, an interaction of cloze X electrode [$F(25,775) = 7.78$, $p < .0001$] explored in further detail through distributional analysis. An interaction of cloze X hemisphere [$F(1,31) = 5.14$, $p = .0305$] was revealed, with ≥50%cloze nouns slightly more positive than <50% cloze nouns (by 0.09μV) over the right hemisphere but showing the opposite pattern over the left hemisphere (low cloze nouns more positive than high cloze nouns by 0.19μV). Cloze X anteriority [$F(3,93) = 12.00$, $pHF = .0001$] also interacted, with a pattern resembling the N400 effect (increased negativity to low cloze continuations) at posterior sites, but the opposite pattern at frontal sites (<50% cloze nouns more positive than ≥50%cloze nouns). This effect was mediated by a higher order interaction of cloze X anteriority X laterality [$F(3,93) = 11.79$, $pHF = .0001$] which indicated that over medial sites there was a strong bias towards an N400-type effect except at prefrontal sites where <50% cloze nouns were more positive than ≥50%cloze nouns: at lateral sites there was virtually no effect of cloze except over the occipital scalp where the N400-like pattern was exhibited.

In sum, in the LP time window there was no main effect of cloze or interaction with VF. However, distributional analyses revealed an N400-patterning effect (low cloze more negative than high cloze nouns) over right hemisphere sites, but not the left hemisphere sites, where low cloze nouns were more positive than high cloze nouns. There was also a reversal from an N400-patterning effect at posterior sites to low cloze nouns being more positive than high cloze nouns at prefrontal sites, particularly over medial scalp regions.

5.5.3.1.2. Cloze analysis within individual constraint level

If cloze probability alone is modifying LP amplitude in the late time window, then we should expect a gradual increase in ERP positivity as cloze decreases. Visual inspection, however, reveals that ERP mean amplitude in the late time window does not follow this pattern for either VF. In **Figure 5.23**, constraint/cloze conditions are sorted in order of mean cloze of each condition. For the RVF/LH, it appears that another factor (most likely constraint) seems to be interacting with cloze. For the LVF/RH, if anything, the overall trend for both time windows seems to be an increasing ERP negativity (rather than positivity) as cloze decreases.

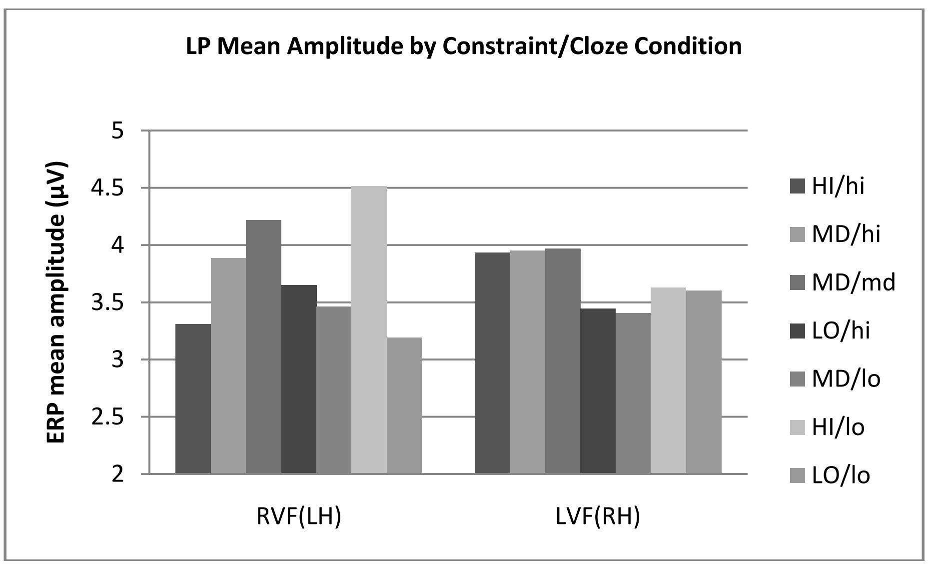


Figure 5.23. Late time window (500-1200 ms). LP mean amplitude within seven constraint/cloze levels in order of each bin's cloze.

Testing the different levels of cloze probability within each level of constraint, we predicted that if there were prediction-related constraint violation effects in the form of a LP, we should observe them to a greater extent for HI constraint compared to MD compared to LO constraint contexts. To test this, we subjected ERPs in the late time window to omnibus ANOVAs at each level of constraint (HI, MD, and LO) within the two visual fields of presentation (LVF and RVF): 2 levels of cloze probability (hi/lo) for HI and LO constraint, 3 levels of cloze (hi, md, lo) for MD constraint X 26 electrodes. See **Figures 5.11,12, and 13** for ERP plots of these conditions. Findings are summarized in **Table 5.8**.

Table 5.8. Late time window (500-1200 ms). Cloze probability findings within constraint levels. Significant and marginally significant findings are highlighted in bold.

		RVF (LH)	LVF (RH)
HI Constraint	Cloze main effect (hi vs. lo)	**F(1,31) = 6.08, p =.0194* (LP-like pattern)**	F(1,31) = .45, p =.5083, n.s.
	Cloze X electrode	**F(25,775) = 2.26, pHF =.0673, marginal**	**F(25,775) = 2.53, pHF =.0627, marginal**
	Distributional effects	n.s. interactions with cloze	**Clz X Lat X Ant, F(3,93) = 12.80, pHF <.0001 (Figure 5.24)**
MD Constraint	Cloze main effect (hi vs. md vs. lo)	F(2,62) = 1.85, p =.1654, n.s.	F(2,62) = 1.44, p =.2436, n.s.
	Cloze X electrode	F(50,1550) = 1.76, pHF =.1083, n.s.	**F(50,1550) = 2.08, pHF =.0386***
	Distributional effects	NA	n.s. interactions with cloze
LO Constraint	Cloze main effect (hi vs. lo)	**F(1,31) = 4.13, pHF =.0508, marginal (N400-like pattern)**	F(1,31) = .11, p =.7475, n.s.
	Cloze X electrode	F(25,775) = .97, pHF =.4252, n.s.	F(25,775) = 1.13, pHF =.3472, n.s.
	Distributional effects	NA	NA

5.5.3.1.2.1. HI constraint contexts only: Effect of hi versus lo cloze across the scalp, over time

As an effect of cloze probability in the late time window was significant primarily for the HI constraint contexts (**Table 5.8**), we wanted to explore this relationship in further detail by analyzing its timing and distribution. To do this, we measured directionality and strength of the ERP effect (the mean amplitude of high cloze subtracted from low cloze continuations) at each electrode site, topographically mapping these differences across time. In time increments of 100 ms, we tracked the cloze effect for each VF beginning in the N400 time window through to the LP time window, spanning 300-1200 ms.

For the LVF(RH) cloze effects (**Figure 5.24**), the topography maps indicate that from 300-500 ms, there are N400 effects of cloze maximal at central posterior scalp sites, where N400s are typically largest. Beginning in the 500-600 ms time window, the cloze effect reverses at prefrontal electrodes, such that HI/lo continuations are more positive than HI/hi continuations at prefrontal sites (an LP-like effect), while posterior scalp sites continue to exhibit more N400-like patterns. The frontal LP-like effect strengthens through the 900-1000 ms time window (maximal LP effect is between 2-2.5 μV at RMPf), after which the effect begins to diminish.

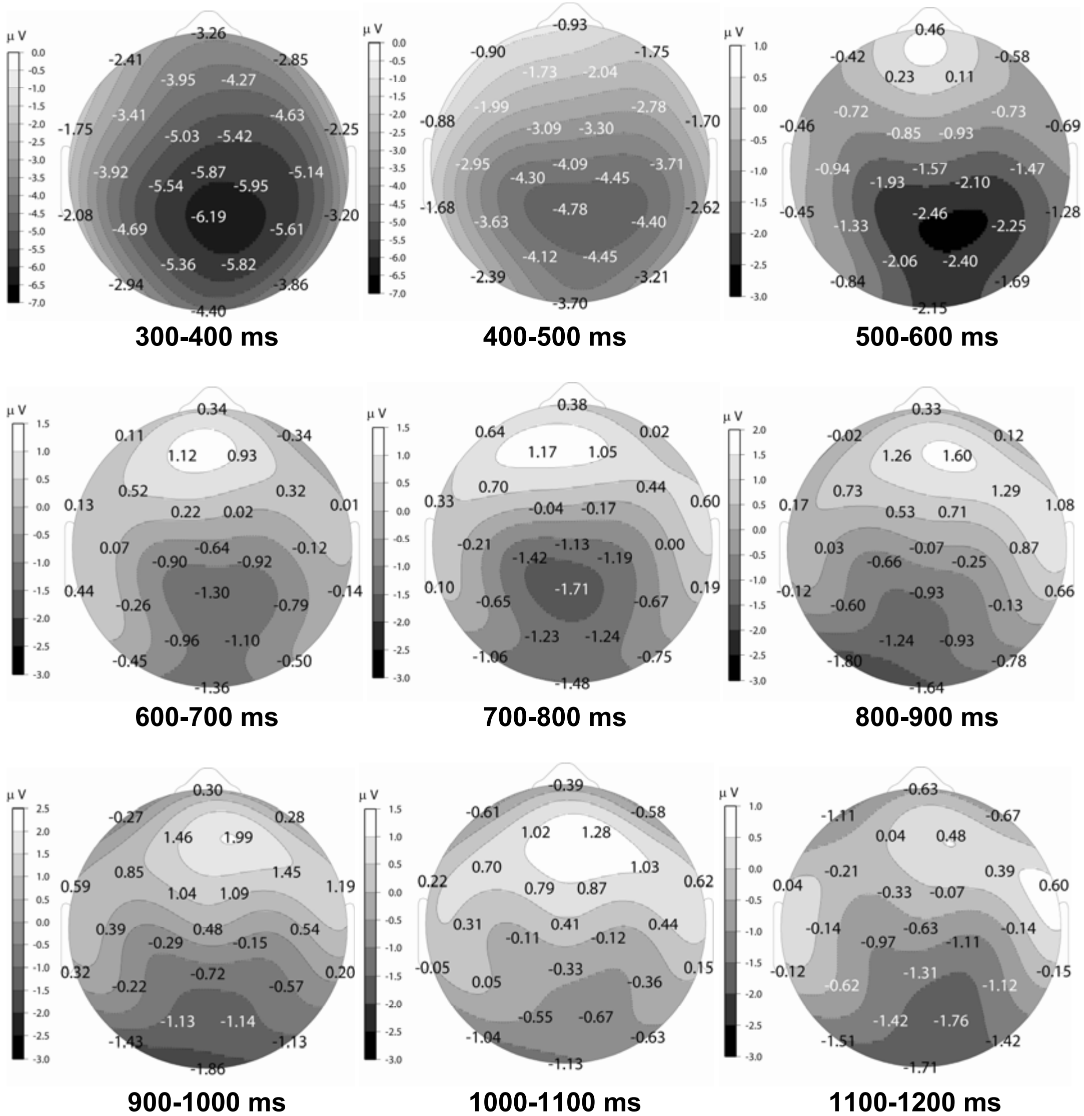


Figure 5.24. LVF(RH) HI constraint contexts, cloze effects (ERP mean amplitude of low minus high cloze continuations) in 100 ms time increments, from 300-1200 ms. Negative ERP differences (darker shading) indicate more N400-like effects of cloze, whereas positive ERP differences (lighter shading) indicate more LP-like effects in which low cloze is more positive than high cloze.

In contrast to the LVF(RH) cloze pattern, for RVF(LH) – shown in **Figure 5.25** – a purely N400-like pattern of cloze is only visible in the 300-400 ms time window. For RVF, the LP-like pattern of cloze begins earlier (beginning already in the 400-500 ms time window at pre-frontal electrodes), gradually becoming widespread over the entire scalp. Similar to

LVF, however, the largest LP-like effects are present at frontal electrode sites, though for RVF the effect size reaches a larger maximum (a condition difference between 3.5-4 μV), which is sustained from 800-1200 ms.

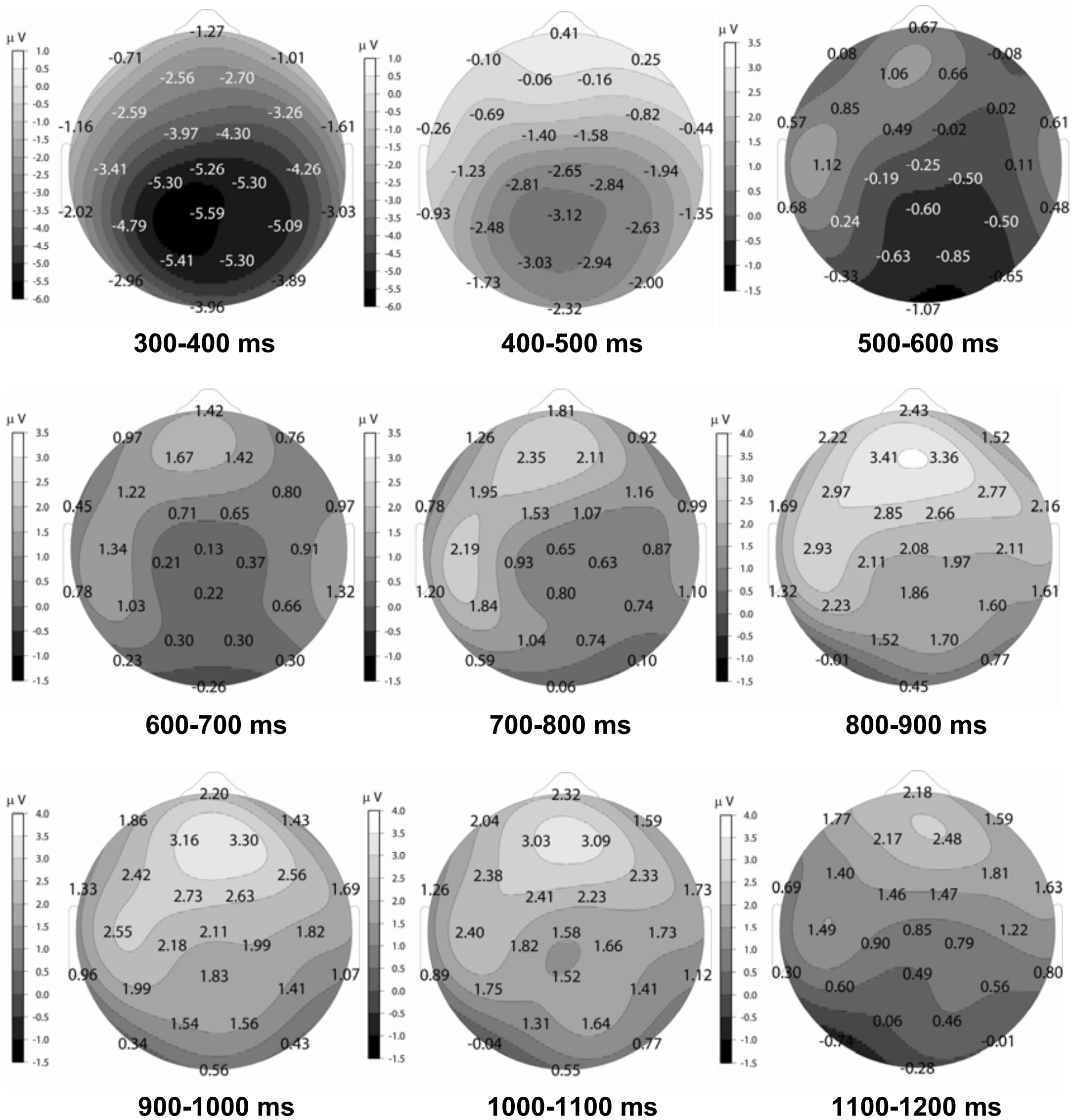


Figure 5.25. RVF(LH) HI constraint contexts, cloze effects (ERP mean amplitude of low minus high cloze continuations) in 100 ms time increments, from 300-1200 ms. Negative ERP differences (darker shading) indicate more N400-like effects of cloze, whereas positive ERP differences (lighter shading) indicate more LP-like effects in which low cloze is more positive than high cloze.

In sum, for the late time window, we predicted larger LP cloze effects within HI constraint contexts and the smallest or no effects within LO constraint contexts. This prediction was upheld. We observed that only at HI levels of constraint were LP cloze effects significant (or marginally significant), while for LO and MD constraint contexts, there was no evidence of an LP effect over either hemisphere. Additionally, we proposed that if Federmeier & Kutas' (1999) proposal of the LH being more predictive and the RH being more integrative holds, then we would expect more pronounced LP effects for RVF presentation. For the HI constraint contexts then, for the RVF(LH) these LP effects were present as early as 400 ms, were maximal at anterior channels (an effect size between 3.5-4 μV), and were widely distributed across the scalp. In contrast, for the LVF(RH), the LP-like effect began later (between 500-600 ms) and was much more focal – though still maximal at anterior channels (an effect size between 2-2.5 μV). This LP-like effect at medial, frontal sites reversed to more N400-like patterns over posterior sites.

5.5.3.1.3. 500-1200 ms time window correlations of noun cloze and ERP mean amplitude

Correlations were calculated between cloze probability and target noun mean amplitude in the late time window for both the LVF and RVF. The correlation mapping is as follows (**Figure 5.26**):

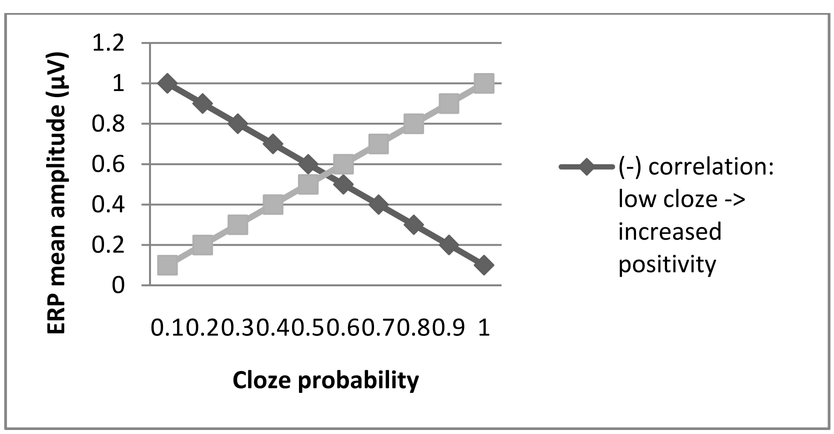


Figure 5.26. Possible correlation patterns between cloze and LP mean amplitude.

The directionality of the late time window mean amplitude/cloze correlation *r*-values is similar for both VFs, but these are non-significant for LVF presentation (**Figures 5.27a, b**). Over anterior scalp sites, the correlation pattern is generally more negative – in other words, the lower the cloze probability of the presented noun, the more positive the mean amplitude of that item (an LP-like pattern). However, this pattern reverses over posterior scalp sites - positive correlation values indicate that the less expected a target is, the more negative the ERP mean amplitude is (a more N400-type correlation).

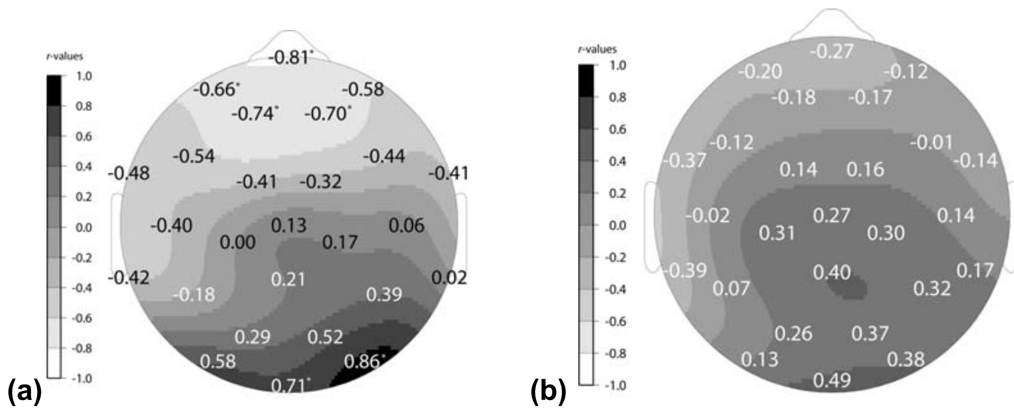


Figure 5.27. Late time window (500-1200 ms). Noun cloze X mean amplitude correlations. (a) RVF(LH) and (b) LVF (RH). Significant correlations (p ≤ .05) are indicated with asterisks (*).

In sum, for the LP time window there was a pattern of increased positivity as cloze probability decreased, represented by higher negative correlations at anterior sites. This pattern reversed over posterior sites, with lo cloze nouns generating more negative ERPs (a more N400-like correlation pattern). This pattern was stronger for RVF than LVF presentation.

5.5.3.2. Effects of traditional constraint violation

5.5.3.2.1. Analyses using 3 levels of constraint with only low cloze conditions

Based on both visual inspection as well as previous studies which have noted an enhanced positivity following the N400 for less expected continuations to high constraint contexts, the ERPs in the late time window (500-1200 ms) were analyzed. This analysis will contrast conditions where the contextual constraint varies, but the cloze probability of the target noun is held constant – in other words, for continuations with low cloze probability in HI, MD, and LO constraint sentences. An omnibus ANOVA with 2 levels of visual field (LVF vs. RVF) X 3 levels constraint/cloze (HI/lo, MD/lo, LO/lo) X 26 electrodes was conducted. This analysis showed no main effect of VF [$F(1,31) = 0.29$, $p = .5939$], nor was there was an overall significant main effect of constraint [$F(2,62) = 2.57$, $pHF = .0864$]. Notably, though, there was a significant interaction of VF with constraint [$F(2,62) = 3.40$, $p = .0398$]. See **Figure 5.28** below. There was also a significant interaction of constraint X electrode [$F(50,1550) = 2.55$, $pHF = .0110$], though this did not interact with VF of presentation [$F(50,1550) = .71$, $pHF = .6785$]. A distributional interaction of constraint X hemisphere was mediated by a higher order constraint X hemisphere X laterality interaction [$F(2,62) = 4.61$, $pHF = .0136$], which indicated that there was a lack of constraint effect (increasing positivity with constraint violation) at left lateral sites (**Figure 5.29**).

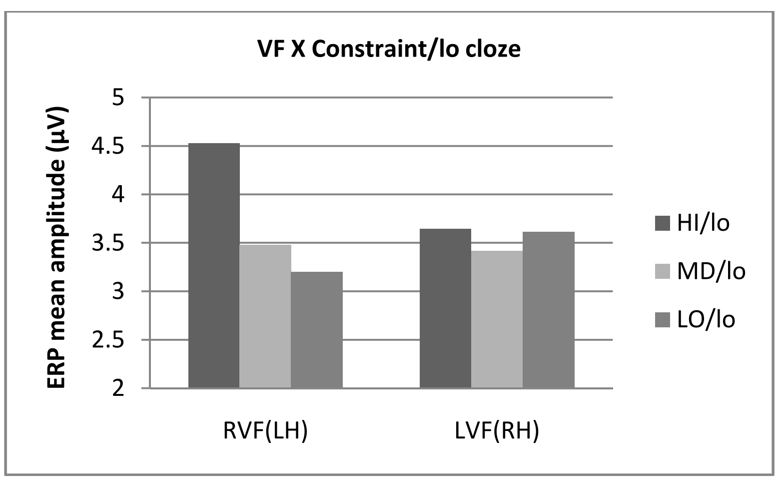


Figure 5.28. 500-1200 ms, VF X constraint/lo cloze.

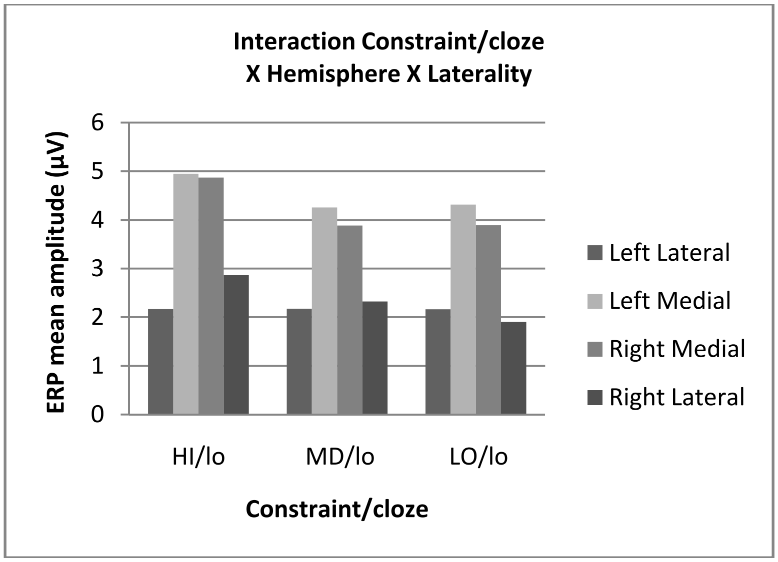


Figure 5.29. LP (500-1200 ms) constraint/low cloze interaction with hemisphere and laterality.

Following up on the interaction of VF with constraint, pair-wise contrasts explored the interaction between VF and constraint. This significant effect was being generated primarily by the RVF Hi/lo condition, which showed an enhanced positivity and contrasted significantly with all other levels.

In sum, between 500-1200 ms, there was no main effect of VF. There was no main effect of constraint, although there was an interaction of VF X constraint which indicated increasing positivity the more constraint was violated for the RVF, but not LVF, presentation. There was a distributional effect of constraint which indicated that the constraint effect diminished at left lateral sites compared to medial and right lateral sites.

5.5.3.3. Correlations of traditional constraint violation

Correlations were also conducted for the low cloze nouns of all three constraint conditions (HI, MD, LO). In a more fine-grained manner, the low cloze sentences were sorted by contextual constraint into 5 approximately 20% bins, and ERP mean amplitude and constraint were correlated for the late time window. In this time window, it was anticipated that if there was a systematic relationship between violation of constraint and ERP positivity in the late time window, then we may observe larger positive correlations (as the sentential constraint of a low cloze item increases – in other words, the degree to which a constraint is violated – the ERP mean amplitude should become more positive). See **Figure 5.30.**

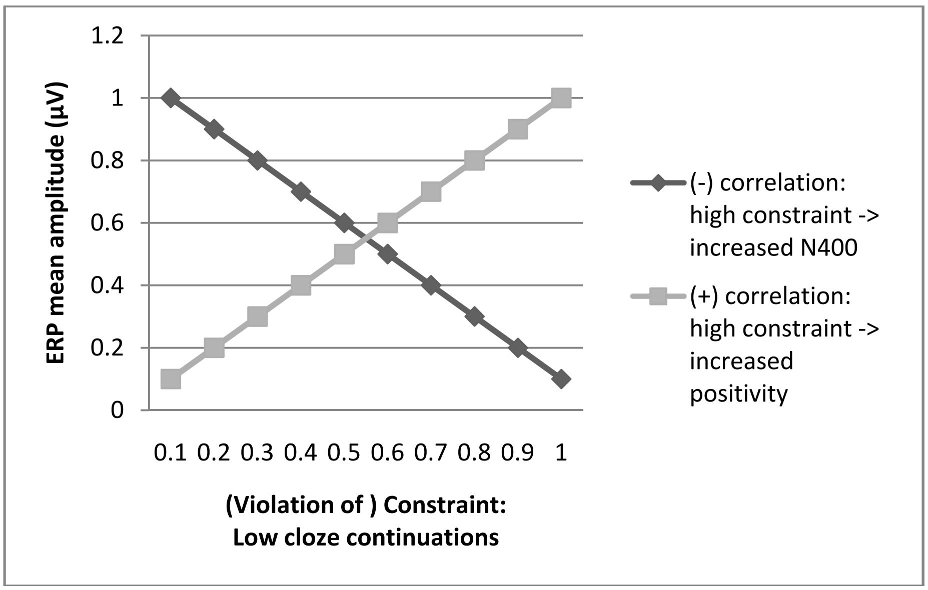


Figure 5.30. Possible correlation patterns between traditional constraint and LP mean amplitude.

This indeed appeared to be the case for low cloze nouns sorted on constraint presented to both visual fields (**Figure 5.31**), with both visual fields exhibiting patterns of widespread, strong positive correlational relationships over the late time window:

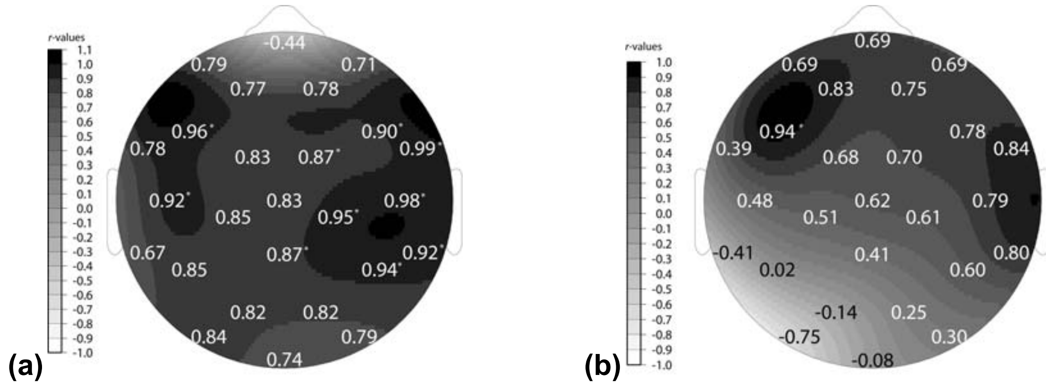


Figure 5.31. LP time window (500-1200 ms). Correlations of mean amplitude with low cloze items sorted on constraint of context (HI+MD+LO/lo items), (a) RVF(LH) and (b) LVF(RH). Significant correlations (p ≤ .05) are indicated with asterisks (*).

In sum, for the LP time window, there was a pattern of increased positivity to low cloze continuations as sentence constraint increased. Though the distribution patterns of high correlation values were widespread for both VFs of presentation, values were more focused over anterior sites for LVF presentation.

5.5.3.4. LP/constraint violation correlations using alternate measure of constraint: effects of recovery strategy

If the positivity in the late time window is associated with sentence constraint – or, better said, the degree to which that constraint is violated – then an alternative measure of constraint would be to determine the strategy norming participants took when the indefinite article (*a* or *an*) in a highly constraining context was substituted by the opposite article type. Depending on the context, norming participants generally supplied an adjective (or adjective phrase) + the expected noun to varying degrees, with the optional strategy being to simply supply a different noun consistent with the presented article type. For each HI constraint sentence context then, we calculated from the norming results the cloze of the expected noun as it appeared *anywhere* in the norming response (no longer as the first word of the response). We then sorted the versions of the HI constraint sentence contexts presented with unexpected articles into ten 10% bins spanning 0-100% according to the percentage of respondents using the noun-salvaging strategy. In each such bin, the mean ERP amplitude, the mean "alternative constraint" value, as well as the cloze probability of the actual noun continuations, were computed (**Table 5.9**).

Table 5.9. Alternative constraint bin breakdown.

CONSTRAINT	% Responses salvaging high cloze noun in version of high constraint sentence with unexpected article	Number of Items in Bin	Avg. cloze of presented noun for high constraint sentences presented with less expected article	cloze
HI	90-100%	12	2%	lo
	80-89%	15	3%	
	70-79%	16	9%	
↓	60-69%	19	5%	↑
	50-59%	24	10%	
	40-49%	17	11%	
LO	30-39%	11	18%	hi
	20-29%	11	18%	
	10-19%	10	20%	
	0-9%	6	62%	

In this analysis, a positive correlation indicates that the more likely that norming participants were to hang onto the high cloze noun by means of an adjective when they received a version of the high constraint sentence with the unexpected article (the higher the alternative constraint was), the more positive-going the ERP mean amplitude is to the actually-presented alternative noun. So for a context where norming results show a large percentage of respondents salvaging the high cloze noun by means of an adjective, by default the actually-presented noun will have a low noun cloze. Conversely, it is the case that for contexts where norming respondents were *less* likely to salvage the high cloze noun (the lower the alternative constraint), the cloze probability of the actually presented noun was higher. (See **Table 5.9.**) So a positive correlation would also indicate that as noun cloze decreases, ERP mean amplitude would become more positive (refer to plots in **Figure 5.32**).

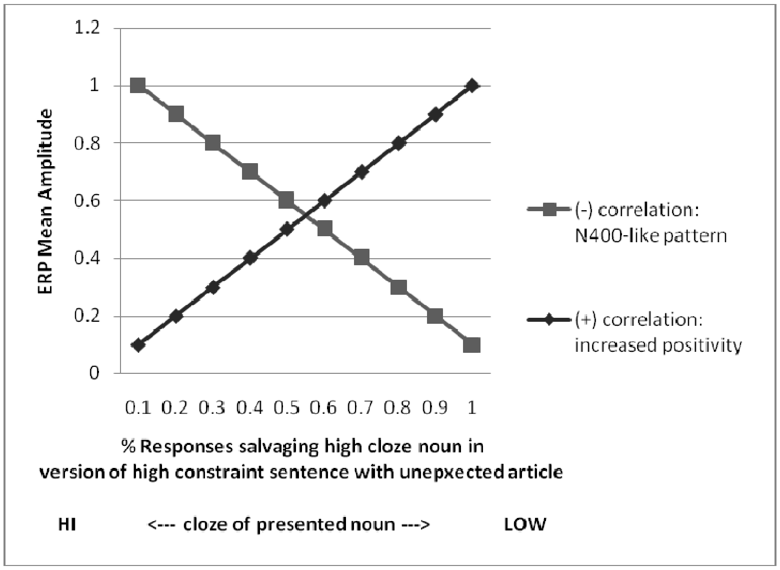


Figure 5.32. Possible correlation patterns of mean amplitude in late time window with "recovery strategy".

On the other hand, a negative correlation in this analysis would indicate that the more likely norming participants were to hang onto the high cloze noun by means of an adjective when they receive a version of the high constraint sentence with the unexpected article and a low cloze noun (the higher the alternative constraint), the greater the negativity when they receive the alternative noun. And in this case, a negative correlation would also indicate that the lower the actually presented noun cloze, the more negative the ERP mean amplitude (a more N400-like pattern as shown in **Figure 5.32**).

So very different conclusions can be drawn depending on the direction (strength, and distribution) of the correlations. Negative correlations would indicate that the ERP pattern in the given time window hinges more on the presented noun's cloze probability. A positive correlation would indicate that ERP amplitude is associated more with the degree of "alternative" constraint violation. Thus, over the N400 time window (300-500 ms) negative correlations would be expected, at least at those scalp sites where N400s are typically observed. Less certain is what would be expected over the late time window.

5.5.3.4.1. LVF(RH)

The correlation values for stimuli presented to the left visual field are presented topographically in **Figure 5.33** below, with r-values for the individual electrode sites, over both the late time window and the earlier N400 time window (300-500 ms):

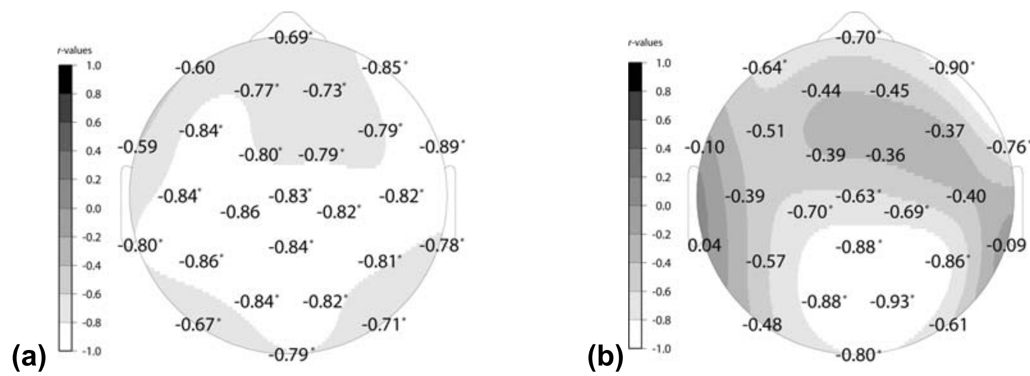


Figure 5.33. LVF(RH) Correlations of mean amplitude with percentage of responses where expected noun is salvaged in version of high constraint sentence presented with unexpected article, (a) 300-500 ms and (b) 500-1200 ms. Significant correlations ($p \leq .05$) are indicated with asterisks (*).

As predicted, in the N400 time window there are strong negative correlations, which are widespread but particularly high over posterior scalp sites where N400 effects are known to be largest. In the late time window, we also observe widespread patterns of negative correlations, however with the highest values concentrated over both posterior sites and right frontal sites. Again, a negative correlation here means that the higher the "alternative" constraint (but also the lower the noun cloze), the more negative the ERP waveform to the actually presented noun.

For the LVF(RH) then, the correlation pattern in the late time window appears to carry over from that observed within the earlier N400 time window (300-500ms) – a pattern more closely indexing the cloze probabilities of the actually presented items, rather than a response to the violation of an expectation.

5.5.3.4.2. RVF(LH)

For the RVF, we observe very different results, with positive correlations that are evident frontally within the N400 time window and which strengthen and become more widespread over the late time window (**Figure 5.34**):

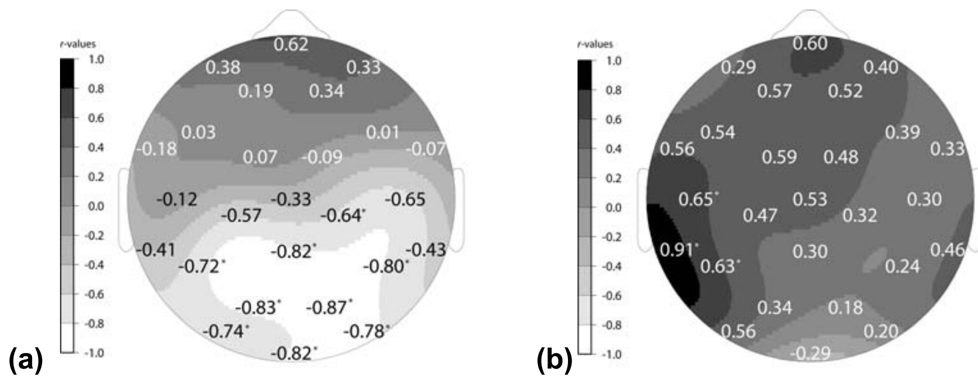


Figure 5.34. RVF(LH) Correlations of mean amplitude with percentage of responses where expected noun is salvaged in version of high constraint sentence presented with unexpected article, (a) 300-500 ms and (b) 500-1200 ms. Significant correlations ($p \leq .05$) are indicated with asterisks (*).

For RVF presentation, during the N400 time window the correlation pattern at posterior sites, as would be expected, seems to be driven more by the relationship of N400 mean amplitude to cloze probability, similar to the pattern observed for the LVF nouns in this 300-500ms time window. However, at frontal sites, we already observe a hint of a reverse in correlation pattern, where positive correlations indicate a sensitivity to "alternative" constraint violation. This is notable not just because it differs from the LVF pattern, but also because a positive correlation is opposite of what would be expected if cloze probability was the main factor modulating ERP mean amplitude. This pattern of positive correlations is present over the late time window, in particular over frontal and left electrode sites.

In sum, for LVF presentation, over the 300-500 ms time window, and continuing into the LP time window, there was a widespread N400-like pattern in which ERP mean amplitude became more negative as noun cloze decreased (but constraint increased). This is very different than the pattern observed for RVF presentation, where in the 300-500 ms time window mean amplitude reflected the typical N400-like sensitivity to cloze at posterior sites, but showed a sensitivity to constraint violation at frontal sites (ERP mean amplitude became more positive with increasing constraint violation). This pattern of increasing ERP positivity with constraint violation was more widely distributed, but still more frontal, over the late time window.

5.5.3.5. Low constraint sentences due to article or lack of context

As should be apparent already from our multiple versions and the varying results of our constraint analyses, the idea of what constitutes a "constraining" context is not unidimensional: constraint could be characterized by a range of different circumstances. For instance the following two sentences (at least in terms of the present experiment) are considered to be low constraint:

Low constraint due to Article:

(1) *Bart did not clean his wound properly. He ended up getting a...*

*...**scar** soon after.* (*an infection* is the expected continuation)

Low constraint due to Lack of Context:

(2) *Valerie didn't know what to make for dinner. At the supermarket she bought a...*

*...**chicken** and a roast.* (many possible acceptable continuations)

In example sentence (1) the context up to the point of the target prenominal indefinite article would be considered highly constraining (if normed without the article). Only at the last minute, when the article before the (lateralized) target noun appears, does the sentence switch from being one of high to being one of low constraint. Instead of subjects converging on the expected noun *infection*, they are forced by the article to come up with something else. In contrast, sentence (2) is low constraint because the context simply does not narrow down the number of possible endings in any significant way.

These differences motivated us to conduct an analysis contrasting these two different types of "low constraint" contexts, the idea being that they may lead to differential positivity patterns in the late time window. More specifically, if the positivity is related in some manner to violation of a highly constrained expectation, one might expect to observe a late positivity in the case of the Low constraint due to Article sentences, but not – or not as much – in the Low constraint due to Lack of Context sentences.

To contrast these two different low constraint conditions, individual Low constraint items (110 contexts) were sorted according to their norming responses. We (somewhat arbitrarily) classified a context to be low constraint due to the "article" if there was ≥20% drop in the context's constraint from when the context was normed with one article type versus the other type (86 items). A low constraint context was deemed as such due to its "context" if there was <20% difference in constraint from norming with one article type versus the next (24 items). To compare these versions of low constraint, we conducted an item analysis within the LP time window (500-1200ms), where an omnibus ANOVA was performed using 2 levels VF X 2 levels of low constraint type (due to Article or Context) X 26 electrodes.

Our analysis revealed a main effect of constraint type [$F(1,108) = 4.15$, $p = .0441$], with continuations to low constraint contexts due to Article showing greater mean amplitude (3.62 μV) than those due to Context (2.69 μV), see **Figure 5.35**. There was *not* a significant interaction of VF x constraint type [$F(1,108) = 1.95$, $p = .1655$, n.s.], though all pair-wise comparisons of the four levels differed significantly from each other (see **Figure 5.36**).

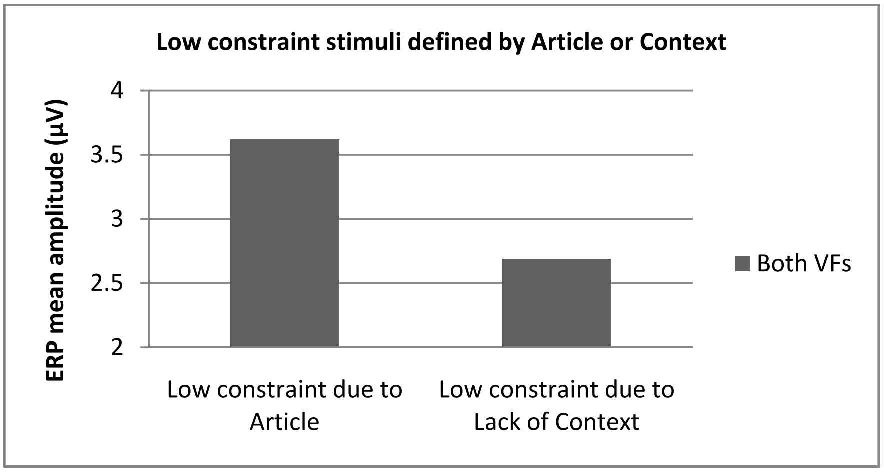


Figure 5.35. Main effect of differently parameterized versions of Low constraint.

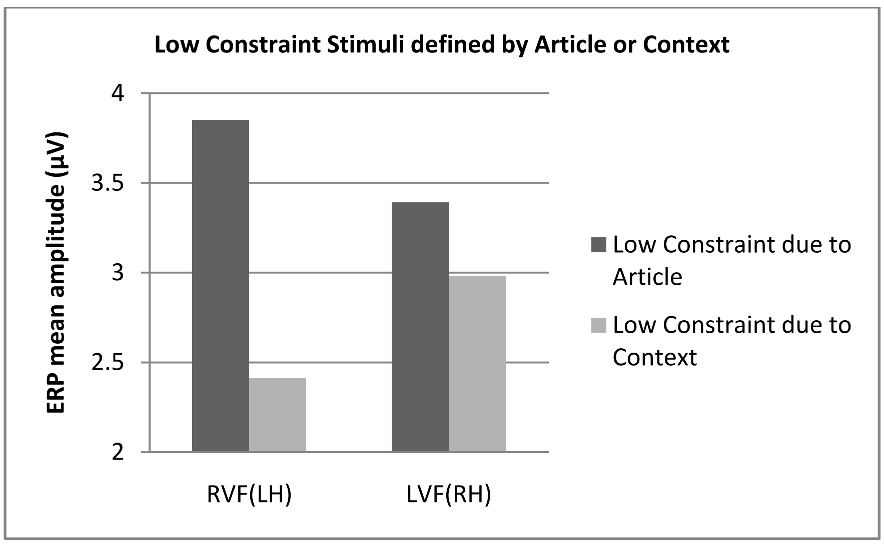


Figure 5.36. Interaction of different versions of Low constraint with VF.

In sum, over the LP window, in contexts deemed Low constraint due to the Article, there was greater positivity than for contexts deemed Low constraint due to Lack of Context. The interactions of LO constraint type with VF were not significant, but there were larger ERP differences between the two LO constraint types for RVF than LVF presentation.

5.5.4. P200 time window (200-300 ms), 11 frontal electrode sites

Per Wlotko & Federmeier (2007), where a P2 effect of constraint was found, in which strongly constrained sentence endings showed larger P2s than weakly constrained endings for RVF (LH) but not LVF (RH) items, we conducted a similar analysis. Using 2 levels of VF X 3 levels of constraint (HI, MD, and LO constraint, collapsed across cloze) X 11 frontal electrode sites (indicated in **Figure 5.37**), we found a main effect of VF [F (1,31) = 12.49, p = .0013] and a main effect of constraint [F (1,31) = 9.87, p = .0002] (**Figure 5.38**), but notably no interaction of VF x constraint, nor VF X constraint X electrode interaction.

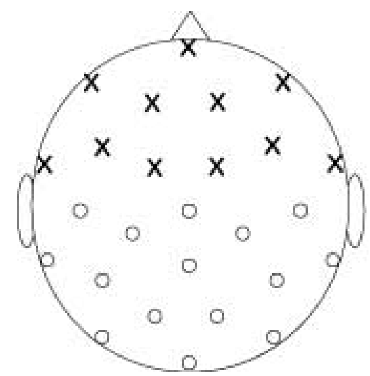


Figure 5.37. Locations of the 11 frontal electrode sites used for N1 and P2 analyses.

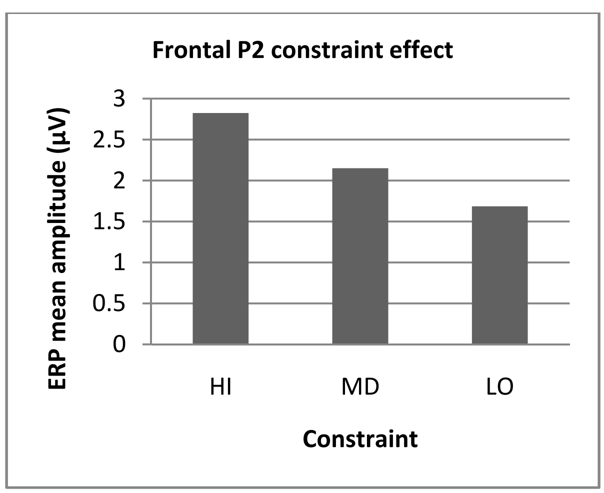


Figure 5.38. P2 (200-300 ms) mean amplitude for HI, MD, and LO constraint (collapsed across cloze) and VFconditions using 11 frontal electrode sites.

We also examined effects of constraint in the P2 (200-300 ms) time window by comparing mean amplitude differences of only the low cloze continuations within the HI, MD, and LO constraint contexts. In this way cloze probability was held relatively constant while constraint was manipulated. We conducted an ANOVA using 2 levels of VF (LVF, RVF) X 3 levels of constraint/cloze (HI/lo, MD/lo, and LO/lo constraint) X 11 frontal electrode sites. There were no main effects of either VF [F (1,31) = 1.55, p = .22, n.s.] or constraint [F (2, 62) = 0.44, pHF = .65, n.s.], nor were there interactions of VF X constraint or VF X constraint X electrode.

In sum, when constraint levels were collapsed across cloze, there was a pattern of increased positivity with increasing constraint, though not differentiated by VF. However, this is somewhat problematic as this pattern of results may also be reflecting the average cloze probability of the constraint conditions, thus possibly indexing a pattern of activity more associated with that in the subsequent N400 time window – the early stages of which are often taken to include the 200-300 ms (our P2) time window. With cloze held constant (using only HI,MD,LO/lo conditions) there was no effect of constraint, either overall or within left or right VFs.

5.5.5. N1 time window (100-200 ms), 11 frontal electrode sites

Similar to the analysis conducted in the P2 time window and based on the effects of contextual constraint on N1 amplitude noted in Experiment 3B of this thesis, we sought to determine if constraint influenced ERP mean amplitude between 100-200 ms.

Using 2 levels of VF X 3 levels of constraint (HI, MD, and LO constraint, collapsed across cloze) X 11 frontal electrode sites, there were no main effects of VF or constraint, nor did the two factors interact.

Like our P2 analysis, we also looked at constraint by comparing only the low cloze nouns for the different constraint levels. Again using 2 levels of VF X 3 levels of constraint/cloze (HI/lo, MD/lo, and LO/lo) X 11 frontal electrode sites, we found no main effects of either VF or constraint, nor any interaction of the two factors.

In sum, there were no effects of constraint on N1 amplitude.

5.6. General discussion

The goal of this study was to assess potential differences in how the two cerebral hemispheres utilize sentence context to construct semantic meaning during on-line language comprehension. Hypotheses derived from previous ERP work by Federmeier & Kutas (1999b; 2002) have postulated that the left hemisphere uses incoming language input in a top-down manner to preactivate upcoming linguistic content, while the right hemisphere relies more heavily on bottom-up processing, waiting until physical stimuli are received to begin integrating representations with preceding context. In their experiments, incongruent items categorically related to highly expected (high cloze probability) sentence endings showed facilitated (reduced) N400 amplitudes relative to incongruent items unrelated to the expected ending, but only for RVF(LH) processing: LVF(RH) N400 amplitudes to the two incongruent item types were similarly large relative to expected endings. Though typically it is difficult to base arguments for predictive processing on N400 patterns occurring at (not prior to) the target words themselves, the similar low cloze and plausibility ratings of the related and unrelated incongruent items ruled out integrative difficulty as an explanation for the RVF(LH) N400 patterns.

Results from Experiments 1, 2 and 3B of this thesis, as well as those from another study (Federmeier, Wlotko, de Ochoa & Kutas, 2007), indicated that at least for centralized presentation there appears to be another ERP component that may be more clearly linked to predictive language processing. In these studies, a late (frontal) positivity was observed to unexpected (low cloze) words continuing highly constraining sentence contexts relative to both expected endings in high constraint contexts and unexpected words in low constraint contexts. Both research groups have interpreted this effect as reflecting a possible "cost" for mispredicting. However in a lateralized version of their original study, Wlotko & Federmeier (2007) reported that the LP effect was absent for the individual VFs of presentation. These results were unexpected because under the view of a predictive LH, an enhanced LP (indexing constraint violation) would have been anticipated for RVF(LH) but not LFV(RH) processing.

In addition to the absence of a LP in their lateralized study, Wlotko & Federmeier (2007) also observed that N400 amplitude to high and low cloze endings in high and low constraint contexts did not vary strictly with cloze probability over either VF of presentation. These results were also unexpected in light of previous elicitation of hemispheric N400 effects as a function of cloze probability (Coulson et al., 2005; Federmeier & Kutas, 1999; Federmeier, Mai & Kutas, 2005). The authors suggested that the widely reported central presentation cloze probability/N400 inverse relationship may in fact reflect the joint contributions of two different hemispheric processing mechanisms. They also proposed that more fine-grained manipulation of stimulus items might indeed lead to the results that they had anticipated.

To this end, the current study manipulated factors of contextual constraint as well as expectancy level (both operationalized in terms of cloze probability ratings) for sentence continuations lateralized to the left or right visual hemifield. With three levels of constraint defined by the cloze probability of the most commonly supplied norming continuation and full ranges of cloze probability within these levels, we were able to dissociate the two factors, which at their upper ends (by definition) are confounded. While the N400 is thought to index the degree to which an item has been semantically facilitated and is known to be highly inversely correlated with cloze probability (Kutas & Hillyard, 1980), the fully graded nature of this relationship has not, to date, been demonstrated within the individual hemispheres. For reasons discussed above, any pattern of N400 effects noted in our study (for either or both hemispheres) could not on its own speak to the prediction question. However, hemispheric N400 effects in relation to our cloze probability and constraint manipulations would be informative on several fronts. First, through both correlations and more categorical analyses, we predicted that N400 amplitude would increase as a function of decreasing cloze over both hemispheres. Though Wlotko & Federmeier's 2007 findings did not pattern in this way and some behavioral data have suggested that the RH is insensitive to message-level constraint, we believed that our more "fleshed out" stimulus set and the temporal sensitivity of the ERP measure might yield the expected N400 results. We proposed that N400 amplitude increasing inversely with cloze over both hemispheres would argue for similar semantic integration across the hemispheres, and would suggest more nuanced usage of message level constraint (as opposed to just lexically-based meaning construction or anomaly detection) for the RH than has previously been described. In addition, as the N400 has been shown to be insensitive to degree of constraint violation (Kutas & Hillyard, 1984), we similarly had no principled reason to anticipate modulation of the N400 as a function of constraint level over either hemisphere. In order to assess this, we analyzed low cloze continuations to all three sentence constraint levels, thus isolating the two factors by holding cloze constant. We kept in mind, however, that effects of the LP might already be influencing ERPs within the N400 time window, as they did in Experiment 3B of this thesis. If this finding held for the present study, then there was potential for a difference in the strength or distribution pattern of the N400 effect between VFs of presentation.

The other question this study addresses is whether there is hemispheric bias for a "misprediction cost" in the form of a post-N400 LP. If the LH indeed uses sentence constraint to preactivate upcoming items, then if it receives an item that violates a strong expectation, the "cost" should be evident only (or to a greater degree) for items presented to the RVF. Experiment 3B, which took advantage of a partial range of constraint values, suggested (at least for centralized presentation) that there is a graded LP effect that is contingent upon constraint violation. Now, with the full range of constraint (and cloze) as well as the lateralized presentation of our present study, we can more reliably evaluate this effect. Results affirming the modulation of the LP by constraint violation would be important because they would be in line with the prediction hypothesis by demonstrating additional (or differential) processing when unexpected words continue strong but not weak constraint contexts. Importantly, such a finding would implicate contextual representations that had already been formed prior to processing the target word, thus reflecting constraint-based expectancy rather than mere bottom up comprehension at the point the target is received.

Finally, previous work suggested that the frontal N1 and P2, typically associated with visuospatial attention and feature detection, may also be modulated by contextual constraint, particularly for RVF processing in the case of the P2. We will begin by discussing our findings relating to these early ERP components, continuing on to the later, more cognitive ERP components.

5.6.1. Early components (Frontal N1 and P2)

5.6.1.1. Frontal P2

Though previous lateralized studies (Federmeier, Mai, Kutas, 2005; Wlotko & Federmeier, 2007) have noted increases in P2 amplitude correlated with increased contextual constraint particularly for RVF/LH processing, we observed no such hemispheric differences in the current study. In Federmeier et al. (2005), it is possible that the observed RVF/LH P2 constraint effects may have reflected early effects of cloze probability, as the “strong constraint” sentences in their study had high cloze endings and “weak constraint” sentences had lower cloze endings. For the Wlotko & Federmeier (2007) study, the authors report a main effect of constraint on the P2 (highly constraining contexts led to increased P2s relative to those for low constraint contexts for the RVF/LH but not LVF/RH), but they did not report a significant interaction of constraint with expectancy (high and low cloze endings). This interaction would be necessary in order to rule in the possibility that it was constraint, and not the possible early onset of cloze on the N400, that was responsible for differences in the adjacent P2 time window. In our study we observed a hemispherically undifferentiated increase in P2 amplitude with increasing constraint (collapsed across cloze), but no overall or hemispheric effects of constraint when cloze was controlled for by examining only low cloze continuations. We are thus inclined to believe that the differences in P2 amplitude that we observed are more consistent with the contributions from the early portion of the N400, whose directionality would result in a similar ERP pattern.

5.6.1.2. Frontal N1

The same analyses as those used to examine the P2 were conducted for the N1, but did not result in any significant findings. Though decreases in N1 amplitude with increased contextual constraint were found for central presentation Experiment 3B of this thesis, no such effects were evident in the current study. For our central presentation findings we argued that the N1 may reflect decreased attention in cases where sentences were highly constrained. As modulations of N1 amplitude have not been widely reported as a function of higher level cognitive factors such as cloze or constraint, our current results are not so surprising. However, there is one possibility worth considering for the present study regarding the lack of an N1 difference between levels of constraint. If the pre-target indefinite articles acted (subconsciously) as cues to the immediately following, highly salient lateralized noun targets, then it is possible that no matter *what* the level of constraint, the visuospatial attentional system may have been primed to allocate more resources toward processing the (anticipated) upcoming lateralized words. In this case, any effects of constraint on N1 amplitude might have been similar across conditions. Further exploration of this idea, however, is beyond the scope of the current study.

5.6.2. N400 (300-500 ms) time window

5.6.2.1. Effects of cloze probability

In general, both the categorical (ANOVA) and continuous (product-moment correlation) analyses revealed that over both VFs of presentation, cloze probability modulated the amplitude of the N400 in a manner similar to that for centralized presentation: as cloze probability for a noun increased, N400 amplitude decreased. This relationship held within both VFs when the target nouns were sorted into two levels of cloze probability (>50%, 50% cloze), when sorted by the high/(medium)/low cloze conditions within levels of constraint (HI, MD or LO), and also when utilizing the full range of cloze (collapsed across constraint) with the correlation analyses. In particular, the noun N400/cloze correlations were similar in strength and distribution to those for central presentation (correlated maximally at $r = 0.9$ over right posterior sites for both VFs), showing for the first time the N400's fully graded sensitivity to cloze within the individual hemispheres.

The only exception to the typical N400/cloze inverse relationship was in the analysis of low constraint contexts presented to the LVF. For these items, the mean amplitudes of LO/hi (mean cloze 11%) and LO/lo (mean cloze 2%) continuations were statistically undifferentiated though the mean amplitudes of the conditions were in the typical direction of the N400 effect (LO/hi less negative than LO/lo). Without having quantified or controlled for factors of lexical association in our stimuli, we do not know whether there were fewer lexical associates in LO constraint contexts, because if this were the case and if, as Faust et al. (1993 and 1995) and Faust & Kravetz (1998) have suggested, the RH capitalizes more on lexical than message level constraint, then this might offer one possible explanation for the hemispheric discrepancy. However, we do not believe this is the case. Federmeier, Mai, & Kutas (2005) compared N400s to plausible endings in high and low constraint contexts, controlling explicitly for lexical associates, and found that for LVF as much as RVF processing there were reductions in N400 amplitude as a function of sentential constraint. Additionally, inspection of our stimulus sentence pairs suggests that when there *are* lexical associates of the target nouns, the majority tend to be in the first, “context” sentence, and not the second RSVP sentence containing the target noun. The temporal delay between a potential “prime” and the target noun would thus be fairly long (on the order of several seconds), thus arguing against word-word priming effects, which are known to be short-lived (Ratcliff & McKoon, 1988).

In total, our results of the N400/cloze analysis argue in favor of both the LH and RH extracting message level meaning to ease integration of lexical items. These findings are also in accord with previous ERP studies indicating that N400 amplitude decreases for expected relative to unexpected endings in highly constraining contexts over both hemispheres (e.g., Federmeier & Kutas, 1999b; Coulson et al., 2005). Interestingly, though, the N400 patterns revealed by the current study’s use of a full range of cloze probability are inconsistent with those of Wlotko & Federmeier (2007), who did not observe the systematic reduction in N400 amplitude with increasing cloze for either VF of presentation. We can only speculate about possible reasons for the differences in N400 findings between the two studies. One possibility may relate to the lateralized targets’ word position within the sentence stimuli: our study used sentence-medial and theirs used sentence-final targets. ERP effects on sentence final words have often been identified as reflecting additional or perhaps slightly different processing than sentence medial targets (e.g., Hagoort, 2003). A more likely possibility may relate to the fine-grained manipulation of cloze and constraint factors used in our study. Our range for these factors allowed us to examine the related ERP effects on a continuum. Yet another explanation may lie in some subtle difference between the types of stimuli and targets used in the two studies; for instance, relatedness of expected and unexpected targets was controlled for in Wlotko & Federmeier's study, but not in ours. We are not sure how this could account for a graded N400 with cloze in our study but not theirs, but in general stimulus factors like these can have important and sometimes unanticipated consequences on resultant data patterns.

5.6.2.2. Effects of constraint

Because factors of constraint and cloze probability are confounded at their upper ends, the most informative way to evaluate effects of constraint is to compare low cloze continuations across constraint levels, thus holding cloze probability constant. In doing this, a 3-way analysis (HI/lo, MD/lo, and LO/lo) unexpectedly revealed that for RVF, but not LVF presentation, ERP amplitude to low cloze items became more *positive* as constraint increased within the N400 time window. Additionally, correlations performed between 300-500 ms confirmed the pattern of increasing positivity with constraint violation. These correlations were strong and widespread for RVF presentation, with maximal correlations present over left anterior scalp sites. Notably, the correlations for LVF processing similarly indicated increasing positivity for constraint violations over anterior sites, though the maximal r-values were weaker and more focal than those for RVF processing. These results are important because factors of constraint (dissociated from cloze) have not previously been found to affect ERP amplitude within the N400 time window. In fact, Kutas & Hillyard (1984) showed that the N400 was *not* modulated as a function of constraint violation, concluding that the component did not reflect a processing “cost”. Our results similarly indicate that N400 amplitude does not increase with the degree of constraint violation; instead, the opposite pattern, one of increased positivity with increased constraint violation, is revealed. Our analyses indicate that this pattern is attributable to the gradual increase in LP amplitude throughout the N400 time window, larger between 400-500 ms than between 300-400 ms. These findings of an LP to constraint violations (stronger for RVF processing) are consistent with the LP results from central presentation for Experiment 3B of this thesis, and like the “early” LP effects from that study, indicate that evidence of a consequence for “misprediction” may already be manifest within the 300-500 ms time window, earlier than the 500-900 ms LP observed by Federmeier, Wlotko, De Ochoa & Kutas (2005) and also earlier than is evident from visual inspection of the waveforms.

The finding of an increased ERP positivity to constraint violations between 300-500 ms strongest at anterior sites primarily for LH processing is notable for a variety of reasons. Perhaps most striking is the hemispheric asymmetry of this effect. Our results are consistent with the proposal of the LH comprehension mechanism functioning more predictively, and thus being stymied when a highly contextually expected item is replaced by an unexpected one. When context is less constraining and expectations for a particular item are minimal, additional processing (whether that processing turns out to be related to inhibition, overriding or reanalysis of an already-formed contextual representation) is not/is minimally required and ERP positivity is present to a lesser degree.

Also relevant is the early onset of this effect. While post-N400 positivities (LPCs, P600s, and constraint-based LPs) have typically been described as having onsets around 500 ms, our effect appears to overlap the N400 time window. If indeed the timing of this (not so) late positivity replicates, it may provide a possible explanation for why such effects have not been frequently observed in other N400 studies. If the N400 and the positivity are concurrent, it could be the case that the positivity may sometimes be "swamped" by strong and widespread N400 effects. Demonstrating a reliably earlier onset than other late positivities, however, could also make it more difficult to argue that our LP effect is related to other late ERP positivities which have been obtained over a vastly different array of experimental manipulations. As a path toward determining the identity (or at least familial relation) of our LP to preceding ones, it seems that subjecting previous studies where LPCs and P600s have been elicited to the visual-hemi field paradigm and determining possible hemispheric biases on these effects might be one route worth exploring.

5.6.3. Late positivity time window (500-1200 ms)

5.6.3.1. Effects of cloze probability

In Experiment 3B of this thesis, we described a double dissociation where N400 amplitude indexes an item's cloze probability but not the degree of constraint violation of the preceding sentence context, and LP amplitude indexes the degree of contextual constraint violation but not semantic integration difficulty (as indexed by cloze). For this reason, we predicted that LP amplitude would not index cloze probability (directly) over either VF, and this is, in essence, what we observed. In general, a comparison of the mean cloze values for all seven constraint/cloze conditions per VF makes clear that cloze probability alone does not modulate LP amplitude. Rather, constraint appeared to be interacting with cloze. For both VFs, a binary analysis (≥50%,<50% cloze) of the ERPs in the late time window indicated that there was a more N400-like association of cloze with mean amplitude over right posterior sites, which was likely carrying over from the 300-500 N400 time window. This pattern reversed over medial, prefrontal sites, to one of increased positivity to low cloze items. This positivity appears to be driven primarily by the ERP response to HI constraint/low cloze continuations for both VFs of presentation.

A comparison over time of the LH and RH cloze effect for only the HI constraint contexts (using 100 ms increments from 300-1200 ms) revealed that the LP effect initiated earlier (between 400-500 ms), was stronger, and was more widely distributed over the scalp for RVF presentation. In contrast, for LVF presentation the cloze effect for HI constraint contexts was initiated later (between 500-600 ms), was smaller in amplitude and was more limited to anterior, medial sites. The earlier onset of the LP effect for HI constraint contexts when presented to the RVF (evident 100 ms earlier than that for LVF presentation) is potentially an intriguing one in that it is consistent with descriptions of RH language processing as generally being slower (e.g., Chiarello, 2003). The same posterior N400/frontal LP pattern was reflected in the correlational analyses, with strong positive and negative *r*-values for RVF but not LVF presentation. In sum, within the LP time window, the LH seems to be more sensitive than the RH to cloze probability, with an increased negativity (prolonged N400) posteriorally to low cloze items, and increased positivity frontally to low cloze items (but with the primary contribution to this effect coming from HI constraint/low cloze items).

5.6.3.2. Effects of constraint

5.6.3.2.1. Traditional constraint analyses

Examining post-N400 LPs using a traditional constraint analysis, between 500-1200 ms, only the LH reflected a main effect of increased positivity with increased constraint violation. Correlation patterns for the two VFs paint a slightly different picture in that both VFs exhibit a strong modulation of LP amplitude with constraint violations, though distributional patterns of the maximal LP correlations were more anterior and less widely distributed for LVF processing. This analysis confirms that, unlike the N400 where the component's amplitude to low cloze continuations is essentially unmodulated by contextual constraint, the LP is highly sensitive to how much an expectation is violated.

In line with Federmeier & Kutas' proposal that LH language comprehension processes are more anticipatory than those for the RH, we observed an LP effect of constraint violation that was more extended and more pronounced for LH processing. However, not predicted under the Federmeier & Kutas proposal was our finding that the RH, too, exhibited some sensitivity to unexpected items in contexts of varying constraint. Distributionally, the RVF(LH) correlations of constraint violation with LP amplitude (maximal *r*-values > .8 over a majority of scalp sites) were strong and widespread, compared to the more frontally distributed high correlations (r-values that generally fell between .7-.9) for the LVF(RH). These results, which strongly argue in favor of anticipatory comprehension, suggest less of a dichotomy between the processing mechanisms of the two cerebral hemispheres, with the RH – at least under some circumstances – engaging along with the LH in additional processing when constraint-based prediction fails.

One possibility to consider regarding the weaker, more focal – but present – RH LP constraint violation correlations, is that the effect occurs well into the epoch, between 500-1200 ms post lateralized target onset. To our knowledge, this is one of the later language ERP components to be examined using the visual hemi-field paradigm. As the lateralized ERP literature attests, there is a clear initial hemispheric processing advantage for the cerebral hemisphere contralateral to the VF of presentation that has been observed at least on early (e.g., visual N1 and P2) up until N400 effects. Unclear, however, is the duration of this processing advantage, and the extent to which information is eventually shared between the hemispheres. It is worth considering that there may be an upper time limit to the processing advantage, after which information initially apprehended by one hemisphere has been shared with the opposite hemisphere.

These considerations could be relevant to our results, because it may be that by sometime between 500-1200 ms into the epoch, the RH has joined with the LH in whatever functional processing is represented by the LP, be it inhibition, reanalysis, revision, etc. On earlier ERP components, one way to assess whether the hemisphere ipsilateral to that of visual presentation has gained access to information through interhemispheric transfer is to assess potential delays in ERP effects between hemispheres. These delays in information transfer across the corpus callosum are thought to be on the order of approximately 10–15 ms (Hoptman & Davidson, 1994). However, such a comparison is difficult with our LP component, as it is a relatively slow wave, is essentially "peakless", and has an onset that overlaps with the N400 component, making it difficult to assess potential timing differences.

On the whole, regardless of whether RH constraint violation effects are a product of callosal transfer or whether they are indicative of similar but independent hemispheric comprehension processes, the LP effect using traditional constraint measures was stronger for LH than RH processing. As we will describe in the upcoming sections, this is a pattern that was additionally upheld by the other constraint analyses we conducted.

5.6.3.2.2. Alternative constraint analyses

Using an alternative measure of constraint (similar to that used in Experiment 3B of this thesis), which took into consideration how norming participants "recovered" when the expected indefinite article was replaced by the unexpected article in the high constraint contexts, very different ERP patterns were noted for LVF and RVF presentation. In this analysis, cloze probability is no longer held constant – as constraint increases, the cloze of the unexpected noun decreases. However, the differences we noted between the two VFs for this analysis are what we found to be most revealing. The LVF shows a distributionally widespread pattern of high correlations indicating that as constraint increases (and as cloze decreases) ERP negativity increases. These more N400-like correlations continue, though weakening some, from the 300-500 ms into the late time window. Conversely for RVF presentation, while the 300-500 ms time window is characterized by N400-like correlations posteriorally and LP-like correlations anteriorally, these patterns give way to an overwhelmingly LP-like correlation pattern in the later time window; in other words, ERP amplitude becomes more positive as constraint violation increases and as cloze of the presented noun decreases.

This alternative constraint analysis, more so than the traditional constraint analysis, draws a sharp contrast between comprehension processes in the two hemispheres. RH processing exhibits a sensitivity to cloze while LH processing patterns more closely with constraint violation. Here, high constraint means that an item is highly expected at some point, though not necessarily immediately, following an unexpected article. Importantly, a context classified as low constraint under a traditional constraint analysis could be considered high constraint under the present analysis. This alternative constraint measure offers a means of discounting the drop in constraint a context might experience due to an unexpected article, perhaps reflecting a more global expectation for upcoming items, rather than constraint being weighted more by how the article forces phonological fit of the upcoming item. This interpretation seems consistent with the results for the LH, where an LP pattern was evident even though the unexpected article downgraded constraint according to the traditional constraint measure, and therefore the potential for constraint violation was lessened. On the other hand, the results for the LVF analysis indicate that the RH appears to take the unexpected article more at "face value", and in turn adjust its expectations accordingly. It is interesting to consider the dissociation in LH and RH patterns for this analysis in terms of (e.g., Gazzaniga's, 2000) descriptions of the LH being the "interpreter" (more prone to abstracting away from the physical stimulus, and more sensitive to the perceived probability of an event occurring, even if not immediately), and the RH as being more "veridical" (updating representations in a way that does not generalize away from the input). Regardless, what this alternative constraint analysis suggests is that using a stimulus set that controls for constraint defined in a very specific way (i.e., not in terms of local fit with an unexpected indefinite article, but rather by more semantically contextual means) might lead to an even more pronounced hemispheric asymmetry in processing of constraint violation.

5.6.3.2.3. Low constraint defined two ways

In our final analysis of constraint violation in the LP time window, we compared low constraint sentences that were classified as such primarily due to the inclusion of an unexpected indefinite article (*Damon was slightly claustrophobic. He preferred climbing stairs over riding in a…*) to low constraint sentences which were simply more open-ended (*Valerie didn't know what to make for dinner. At the supermarket she bought a…*). The targets in sentences of the first type showed a greater LP than those in the open-ended sentences, and while the hemispheric difference for this effect was not significant, the ERP difference between the two context types was larger for LH processing.

The results of this particular analysis are interesting for two reasons, the first being that there actually *was* any ERP difference at all between noun targets in the two low constraint sentence types. This brings up a general concern with a term like "constraint". Though for the most part "constraint" is considered to be operationalized in a widely agreed-upon way, this "way" is actually not very explanatory. Many dimensions (e.g., syntax, word frequency, concreteness, phonology, world knowledge, etc.) could contribute to a particular word being "constrained", though generally the influence of individual factors is not specified. For instance, in previous norming data (described in detail in Experiment 3A) we have observed that for truncated sentences such as, "*Because they were playing baseball so close to the house, the children ended up shattering an…*", the presentation of the indefinite article drastically alters the offline "constraint" value of the context from when normed with the expected article, *a*. With the expected article, *a*, there is a single, highly expected continuation (*window*) with 96% cloze probability, which by definition means that the constraint value is also 96%. With the opposite article (*an*), the fan of responses is much wider (e.g., [*old, expensive, oval, elegant, open, upstairs, antique*] *window*), and calculating cloze probability/constraint based on the most common response – in this example, *old window* – the constraint is reduced to 27%. However, if constraint is considered in a different way (e.g., the most common word found anywhere within a norming response) the value for this particular example only experiences a slight reduction, to 92% cloze/constraint (for *window*). It is clear, then, that in an ideal world, “constraint” would be considered on a variety of different levels, with, for example, the cloze value for a particular part of speech, a specific phonological pattern, or a set of shared semantic features being calculated separately from norming results. In addition to these concerns, there is also the issue of how “constraint”, as operationalized through cloze norming, underestimates the probability that certain items have a higher likelihood of appearing than their cloze values might reflect. This is particularly true of low cloze items, which though they may not be supplied when testing a limited number of norming respondents, might occur in natural language with a level of frequency higher than, say, implausible continuations.

In particular, though, for our study, the greater LP to targets in low constraint sentences due to the indefinite article is more consistent with the idea that the LP is elicited in instances where sentence content has narrowed the possible continuations considerably, possibly preactivating likely upcoming items. When high constraint is downgraded to low constraint because of a pre-target indefinite article, it seems unlikely that during online processing a highly preactivated item would immediately be dismissed as a possible continuation – either because there is still the possibility that the expected item may be received (following an adjective) or because the brain’s parser takes a “wait-and-see” approach. In either case, though sentences of this type are classified as low constraint, it is clear that they are of a very different variety than the more open-ended low constraint sentences.

The VF difference in this analysis, though not statistically reliable, offers a second point of interest. Again, we see that the effect of increased LP to low cloze continuations in contexts where (up until the pre-target word) a different continuation was highly constrained appears more dominant for LH than RH processing. Our analysis comparing two different types of "low constraint" contexts warrants further, less opportunistic exploration than what we were able to engage in here. However, the fact that these results are consistent with the other hemispheric analyses of the LP in this experiment lends support to the idea of a cost for preactivation that is more lateralized to LH processing.

5.6.3.2.4. Summary of LP constraint effects

Across the variety of LP/contextual constraint analyses we conducted, our results converge on a common pattern of findings. The amplitude of the LP increases with constraint violation, particularly with presentation to the RVF. Our study is thus the first to show that there is a cost to preactivating but not receiving upcoming linguistic input, which is biased towards LH processing. Much like the graded prediction effects observed in Experiment 1, our LP also seems to reflect graded constraint violation processing. While the LP effect is most evident for sentence contexts in which there is a strong expectation for a particular word, the results from correlation analyses indicate that these effects may also extend to less constraining contexts, for which there may not be a single, highly expected continuation. Additionally, while the effect of increasing LP with constraint violation was, in general, stronger for RVF(LH) processing, this pattern was not nonexistent for LVF (RH) presentation. For correlation analyses using a traditional measure of constraint, both VFs showed increasing LPs with constraint violation, albeit with less widespread/more anterior scalp distributions for LVF than RVF presentation. Conversely, conducting correlation analyses using an alternative measure of constraint resulted in a complete absence of an LP effect for LVF processing, where one was present for RVF processing. The obvious explanation for this discrepancy relates to how "low constraint" is operationalized, and certainly future explorations of sentential constraint should take such factors into consideration. However, even more importantly, we believe that the difference in the results of the two constraint analyses may offer a window into how the hemispheres utilize language input to shape contextual representations.

5.6.4. LH vs. RH sentence-based meaning integration and constraint violation

5.6.4.1. Integration processes

We will now elaborate on the two main results of the present study, beginning with our N400 findings relating to semantic integration. In this experiment, graded N400 effects as a function of sentence cloze were undifferentiated in strength or timing by VF of presentation. Importantly, these results suggest that both hemispheres incrementally construct representations of the accruing context and experience similar ease or difficulty in integrating more or less expected words into this representation. These patterns seem to argue against models proposing that semantic activation is always broader and weaker for the RH, but more focal for the LH (Beeman, Friedman, Grafman & Perez, 1994), as both hemispheres benefited from fine-grained constraint provided through facilitative contexts. We have also argued that the facilitative effects of high constraint contexts observed for both hemispheres seem to emanate from message-level, rather than lexical-level, constraint. This is relevant because of theories suggesting that the RH relies mainly on lexical level priming (e.g., Chiarello, Liu, & Faust, 2001; Faust, 1998). It seems unlikely (though admittedly not impossible) that both hemispheres would arrive at such similar N400 patterns through different mechanisms, though based on our stimulus constructions we doubt that lexical-based priming could sufficiently explain such effects. Interestingly, our pattern of hemispheric N400 results also differed from those of Wlotko & Federmeier (2007), who used a similar experimental manipulation. In their study, N400 amplitude was not modulated strictly by cloze for either VF/hemisphere, however when the N400s for the two hemispheres were averaged, the pattern was consistent with that for central presentation N400/cloze patterns. They suggested that neither hemisphere may actually be sensitive to cloze in a graded fashion, and that the canonical cloze-graded N400 may be reflecting the hemispheres' joint influence. Our results argue against this, as N400 amplitude was highly inversely correlated with cloze for both VFs, in both cases over the same scalp areas where N400 effects have been shown to be maximal (right central, posterior sites).

In accordance with previous findings, the N400s to less expected continuations in our study were not modulated by the degree to which a particular context constrained for a different item. Even though it may seem somewhat counterintuitive, N400 amplitude does not reflect a cost for "getting it wrong"; rather, it reflects the benefit of "getting it right" – either wholly or partially. Under an anticipatory comprehension model like the one we have proposed in this thesis, decreases in N400 amplitude may indeed turn out to reflect, at least in part, the fact that the received item or features of that item have been pre-activated by the preceding context. Although in the current study it is impossible to differentiate between predictive and integrative strategies based on the N400s to target nouns alone, both the Experiment 1 and the Federmeier & Kutas (1999) N400 reductions point to the component reflecting at least some degree of preactivation. In turn, the similarities in N400 patterns across hemispheres in the current study would suggest then that *both* hemispheres may be capable of constraint-based prediction, which would argue against a strictly LH predictive/RH integrative division.

One question highlighted by our hemispheric findings, along with the preceding studies of this thesis, relates to the distinction drawn between integrative and predictive processing. Some might argue that any proposed mechanism for eased integration must necessarily involve some kind of preactivation, whether relatively early (e.g., well before an item has been encountered) or later (e.g., temporally closer to target onset). Such suppositions would suggest, then, that the distinction between integration and prediction is an artificial one. However, in terms of what we take preactivation to mean, we would disagree. Though innumerable psycholinguistic findings have shown clear benefits to processing linguistic items within facilitative contexts (reflected in such measures as speeded lexical decision or naming times, reduced N400s, earlier and more frequent gaze fixations, etc.), it does not necessarily follow that targets or their features were preactivated. We take facilitated integration to mean that the language processor takes advantage of some established associations in long term memory (e.g., schema-, frequency-, semantically-, categorically-based) between a target and contextual elements, which allows for eased processing, in a bottom-up fashion, once a target has been received. On the other hand, we suggest that preactivation can involve multiple items or their features being "triggered" in a graded, probabilistic way during the process of online meaning construction, through associations in semantic memory. Such associations could be at levels of familiar meaning relations (e.g., those commonly tested by psycholinguists – e.g., category relations, semantic similarity, etc.), but also through more subtle associations, for instance along the lines of what Gibson (1977) has referred to as "affordances". The best way to make the prediction argument is to find evidence for preactivation prior to a target (as Experiment 1 of this thesis and other studies – e.g., Wicha, Moreno & Kutas, 2003a and 2004; Wicha, Bates, Moreno & Kutas, 2003b; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005 – have done), though in general, such findings are difficult to come by.

So perhaps, then, the difference between the two hemispheres is not that one is predictive and the other integrative, but rather that the two make maximal use of different aspects of the input to preactivate upcoming continuations to different extents. In some cases the contextual information may be specific or strong enough to narrow the upcoming possibilities to a considerable degree for both hemispheres, and in other cases not as much. Determining which aspects of constraint the hemispheres are sensitive to (e.g., category-based information played an important role in Federmeier & Kutas' results) and the potentially differential goals for which the two hemispheres tap into an identical input source (e.g., perhaps the LH language comprehension is more closely linked to the language production specialization of that hemisphere, yielding a more prediction-based strategy, Pickering & Garrod, 2006) may be approaches worth pursuing.

It could be said that prediction turns out to be a useful strategy except in cases where it is not. Though there is a great deal of debate about how predictable everyday language actually is (some would argue prediction occurs only rare circumstances or not at all, e.g., Mitchell, 1982; Jackendoff, 2002; Morris, 2006), even those who believe in anticipatory language comprehension would agree that a predictive system will undoubtedly fail at times. Life, language and everything in between would certainly be very boring if it did not! However our assumption about prediction in language comprehension is that it is just one example of the more general way in which the brain functions, with experience-based prediction being a sort of "default" cognitive principle. As Kveragaa, Ghumana and Bar (2007) put it,

> ...a fundamental function of the brain is to predict proximate events, which facilitates interactions with external stimuli, conserves effort, and ultimately increases the chances of survival.

Nonetheless, if one were designing a language comprehension system, it would be advisable to equip the system with a flexible and fluid way of recovering from prediction-gone-wrong, or with a way of keeping open alternative possibilities along the way. So instead of proposing that unlike the LH, the RH does not processes linguistic material anticipatorily at any level, perhaps RH comprehension also takes place in an anticipatory fashion, but is simply not "weighted" in the same ways. For instance, in line with Beeman's coarse coding proposal, preactivations could be less targeted – broader or diffuse – in cases where there is less constraint, but when constraint is high and there are fewer alternative interpretations for the parser to pursue, the RH, like the LH, can "get it right". Rather than the differences in hemispheric comprehension processes being ones relating to an anticipatory versus wait-and-see approach, perhaps the differences are more ones of degree. Imaging studies of adult LH brain damage patients who recover from aphasia and show increased processing for RH homologous areas suggest that there is at least some potential for the RH to "pick up the slack" for language tasks normally performed by the LH (e.g., Blasi, Young, Tansy, Petersen, Snyder & Corbetta, 2002; Blank, Bird, Turkheimer & Wise, 2003). These studies argue, though, in only a broad way for some overlap in hemispheric language function, and not the mechanism by which the same outcomes are ultimately achieved by the LH and RH.

5.6.4.2. Constraint violation

Although we argue above that predictive language processing may extend to the RH as well as the LH, the LP results based on the stimuli used in our study clearly suggest that there is at least a strong bias for the processing of constraint violations by the LH. A strong test of hemispheric prediction and an experiment that we plan to conduct, is to replicate the current study, except with the prenominal indefinite articles lateralized instead of the target nouns. In this way we can directly compare prediction effects we observed at the article for central presentation (Experiment 1) to those we might observe for the same stimuli processed by the individual hemispheres. Of primary interest would be whether we would observe N400 effects to RVF articles only, or for both VFs. At any rate, the current study indicates that for the cognitive processes that we have related to "cost" of mispredicting, the LH generally is more sensitive. These effects were present over several different analyses which operationalized constraint violation in different ways.

Another point worth highlighting is that what we have generically labeled a "late positivity" actually manifested earlier in the ERP epoch than other such LPs described in the sentence processing literature (e.g., the P600, Syntactic Positive Shift, and LPC). Our effect of increased ERP positivity to constraint violations began as early as the 300-500 ms (N400) time window. So it seems that processes related to semantic integration (as indexed by the N400) and those related to prediction violation may overlap in time. Perhaps, then, for constraining contexts where an upcoming item has experienced strong preactivation, the functional processing relating to the LP does not hinge on a failed attempt at integration. It may be that quite early in the epoch the brain is sensitive to the fact that it did not receive what it expected and initiates whatever additional processing is indexed by the LP (e.g., inhibition, reanalysis, revision, etc.) during the same time period when the N400 peaks. The pattern of P2 findings in our study makes it difficult to conclude whether this is the case, as the only increase of P2 amplitude observed was for an analysis where constraint was confounded with cloze. In this case, we were inclined to interpret the P2 amplitude increases to high constraint (high cloze) items that were present for both VFs as reflecting the onset of the N400 effect, which would pattern in a parallel direction. In sum, this problem of overlapping components makes it difficult to assess the precise onset of the "LP" effect.

The timing and duration (as well as the scalp distribution) of our LP effect are relevant if we want to compare our results to other language studies which have noted similar positivities. In Experiment 3B we attempted to do this for the findings relating to our centralized presentation LP. However, it is difficult to make such comparisons for hemispheric LP effects, as there have been so few language comprehension ERP studies that have utilized the hemifield paradigm to test various theories relating to contextual constraint. We make note, though, of the fact that the LPs obtained in Experiments 1, 2, and 3B of this thesis were highly similar to the hemispheric LP effect observed in the current experiment. The centralized LP thus seems to reflect the dominant contribution of LH constraint processing. It would be informative to test stimuli where "semantic P600s" have been observed (e.g., Kuperberg, Sitnikova, Caplan, Holcomb, 2003; Kim & Osterhout, 2005; Hoeks, Stowe, Doedens, 2004; Kolk, Chwilla, van Herten & Oor, 2003, to name just a few) using lateralized presentation, to determine whether such effects would show a LH bias similar to ours. Though their various theories of the "semantic P600s" do not, to our knowledge, make predictions about the effect being more lateralized to one hemisphere or the other, if there turned out to be some LH bias for their effects, it could be a first step toward developing a more unitary theory about the functional significance of such ERP effects.

In sum, the LH bias for ERP constraint violation effects to unexpected continuations in high but not low constraint contexts held across three different types of analysis: whether constraint was quantified in terms of (a) cloze of the most provided continuation in a norming study, (b) the tendency for a high cloze noun to remain high cloze even in the face of an unexpected article, or (c) by contrasting low constraint contexts that are deemed so either by virtue of an unexpected article (more like high constraint contexts) or because they are simply more contextually open-ended. On this basis, we agree with part of Federmeier and colleagues' proposal about the differential roles of the hemispheres in language comprehension – that is, that the LH preactivates linguistic information prior to its receipt, and engages in additional processing when the predictive process is thwarted. However, our results are not directly in line with the second half the Federmeier proposal, that RH comprehension is purely bottom-up, or integrative, in nature. While the RH in general showed much weaker effects of constraint violation compared to those of the LH, there nonetheless was a pattern of increased post-N400 ERP positivity in the (a) and (c) constraint analyses described above.

The complete lack of a LVF LP effect for the alternative (b) constraint analysis, though, is intriguing. This analysis more than the other two takes into consideration how strongly comprehenders want to maintain their noun expectations, even in the face of an unexpected article. And unlike analyses (a) and (c), results from the (b) analysis indicate that there is a very clear distinction between a predictive LH (with a LP effect in the late time window) and an integrative RH (with an N400-like correlation pattern extending to the late time window). Opposite of the typical N400 pattern, for the LH, the ERP amplitude becomes more *positive* with decreasing cloze (and increasing alternative constraint violation). The RH, on the other hand, shows increasing ERP negativity as cloze decreases. It seems, then, that the LH is able to project the most likely way that a high constraint sentence might continue in the face of even a small bit of prediction inconsistent information, i.e. an unexpected article – either by means of an adjective to salvage the original noun expectation or by an alternative phonologically fitting noun. So for the LH, an unexpected article does not necessarily alter the preactivation level for a highly expected noun, and in a sense, a high constraint context can remain high, even when the unexpected article is presented. However, the RH does not seem to adapt in the same way. Once an unexpected article continues a high constraint context, this information can down-regulate the constraint level of the context, and the ERP to the subsequently presented noun seems to more closely reflect the lowered constraint and integration difficulty for the upcoming noun. These patterns are consistent with the idea that under conditions of limited constraint, the RH may not be as successful as the LH at preactivating upcoming linguistic material. The linguistic sensitivities behind this hemispheric processing difference will undoubtedly require further exploration. In sum, it seems that the RH, like the LH, is capable of preactivating upcoming language information, but only under conditions in which the sentential input provides sufficient constraint. While both hemispheres appear capable of exhibiting fine-grained sensitivity to contextual information, it seems likely that these sensitivities are weighted by different factors or possibly for different cognitive goals.

5.7. Conclusions

In sum, the following main points can be made from the present study. First, regardless of visual field/cerebral hemisphere, the N400 is similarly sensitive to cloze probability as an index of offline expectancy. This finding argues against the proposal that the N400/cloze relationship typically observed for central presentation is an amalgam of the two hemispheres. It also challenges the idea that the RH is insensitive to message-level constraint, as RH N400 amplitude was modified by subtle shifts in constraint values in a manner similar to the LH. Second, the ERP amplitude of the late positivity (LP) was largest and thus most sensitive to strong violations of contextual constraint – i.e., to low cloze continuations of high constraint context – though clearly this effect was a graded one. The LP effect, though revealed over both the LH and RH, was generally stronger and more widely distributed for RVF/LH presentation. Multiple analyses revealed this LH bias, evidence which we have attempted to link to a consequence of mispredicting. Finally, a contrast in the results of analyses operationalizing constraint in two different ways may offer a window into how the two hemispheres differentially use sentence context information to modulate constraint: perhaps with the LH taking a more probabilistic approach and the RH processing language input more veridically, though this proposal requires further exploration. More broadly, the discrepancies in the results of the two constraint analyses should serve as a kind of warning to those interested in studying such aspects of language processing:

considering experimental factors such as “constraint” - or “cloze probability”, for that matter - as unitary concepts and quantifying them as such may obscure important points of variability between stimulus items. While it may be methodologically impractical, if not impossible, to account for all the different features that contribute, e.g., to sentential constraint, it is wise to keep this caveat in mind.

5.8. References

Banich, M.T. (2003). Interaction between the hemispheres and its implications for the processing capacity of the brain. In K. Hugdahl, & R. J. Davidson (Eds.), *The asymmetrical brain*. (261-302). Cambridge, MA, US: MIT Press.

Bates, E., Bretherton, I. & Snyder, L. (1988). *From First Words To Grammar*. Cambridge University Press, Cambridge.

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In M. Beeman, & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. (255-284). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Beeman, M., Friedman, R.B., Grafman, J. & Perez, E. (1994). Summation priming and coarse semantic coding in the right hemisphere. *Journal of Cognitive Neuroscience*, 6(1), 26-45.

Blank, S.C., Bird, H., Turkheimer, F. & Wise, R.J.S. (2003). Speech production after stroke: The role of the right pars opercularis. *Annals of Neurology*, 54(3), 310-320.

Blasi, V., Young, A.C., Tansy, A.P., Petersen, S.E., Snyder, A.Z. & Corbetta, M. (2002). Word retrieval learning modulates right frontal cortex in patients with left frontal damage. *Neuron*, 36(1), 159-170.

Burgess, C. & Simpson, G.B. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*, 33(1), 86-103.

Chiarello, C. (2000). Inferring the nature of semantic processes by varying priming procedure: A reply to Koivisto and Laine. *Laterality: Asymmetries of Body, Brain and Cognition*, 5(1), 23-27.

Chiarello, C. (2003). Parallel systems for processing language: Hemispheric complementarity in the normal brain. In M. T. Banich, & M. Mack (Eds.), *Mind, brain, and language: Multidisciplinary perspectives*. (229-247). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Chiarello, C., Liu, S. & Faust, M. (2001). Bihemispheric sensitivity to sentence anomaly. *Neuropsychologia*, 39(13), 1451-1463.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A, 497-508.

Coulson, S., Federmeier, K.D., Van Petten, C. & Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1), 129-147.

Deacon, D., Grose-Fifer, J., Yang, C., Stanick, V., Hewitt, S. & Dynowska, A. (2004). Evidence for a new conceptualization of semantic representation in the left and right cerebral hemispheres. *Cortex*, 40(3), 467-478.

Faust, M. (1998). Obtaining evidence of language comprehension from sentence priming. In M. Beeman, & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. (161-185). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Faust, M. & Kravetz, S. (1998). Levels of sentence constraint and lexical decision in the two hemispheres. *Brain and Language*, 62(2), 149-162.

Faust, M., Babkoff, H. & Kravetz, S. (1995). Linguistic processes in the two cerebral hemispheres: Implications for modularity vs interactionism. *Journal of Clinical and Experimental Neuropsychology*, 17(2), 171-192.

Faust, M., Kravetz, S. & Babkoff, H. (1993). Hemisphericity and top-down processing of language. *Brain and Language*, 44(1), 1-18.

Federmeier, K.D. & Kutas, M. (1999a). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Federmeier, K.D. & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, 8(3), 373-392.

Federmeier, K.D. & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, 40(7), 730-747.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505.

Federmeier, K.D., Mai, H. & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition*, 33(5), 871-886.

Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E. & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research. Special Issue: Mysteries of Meaning*, 1146, 75-84.

Galuske, R.A.W., Schlote, W., Bratzke, H. & Singer, W. (2000). Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science*, 289(5486), 1946-1949.

Gazzaniga, M.S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain*, 123(7), 1293-1326.

Gibson, J. (1977). The Theory of Affordances. In R. E. Shaw and J. Bransford, Eds., *Perceiving, Acting, and Knowing*. Hillsdale: Lawrence Erlbaum Assoc.

Hoeks, J.C.J., Stowe, L.A. & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19(1), 59-73.

Hoptman, M.J. & Davidson, R.J. (1994). How and why do the two cerebral hemispheres interact? *Psychological Bulletin*, 116(2), 195-219.

Jackendoff, R. (2003). Foundations of language-brain, meaning, grammar, evolution. *Applied Cognitive Psychology*, 17(1), 121-122.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9(11), 712-718.

Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205-225.

Kolk, H.H.J., Chwilla, D.J., van Herten, M. & Oor, P.J.W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1-36.

Kucera, H. & Francis, W.N. (1967). *Computational analysis of present-day English*. Providence, RI: Brown University Press.

Kuperberg, G.R., Sitnikova, T., Caplan, D. & Holcomb, P.J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17(1), 117-129.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

Kveraga, K., Ghuman, A.S. & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, 65(2), 145-168.

Medler, D.A. & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/.

Metcalfe, J., Funnell, M. & Gazzaniga, M.S. (1995). Right-hemisphere memory superiority: Studies of a split-brain patient. *Psychological Science*, 6(3), 157-164.

Mitchell, D.C. (1982). *The process of reading: a cognitive analysis of fluent reading and learning to read*. Published Chichester; New York: Wiley, 1982.

Morris, R.K. (2006) Lexical processing and sentence context effects. In M.J. Traxler & M.A. Gernsbacher (Eds) *Handbook of Psycholinguistics* (2nd Ed.). Academic Press: London.

Neville, H.J., Kutas, M. & Schmidt, A. (1982). Event-related potential studies of cerebral specialization during reading: I. studies of normal adults. *Brain and Language*, 16(2), 300-315.

Ratcliff, R. & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95(3), 385-408.

Scheibel, A.B., Fried, I., Paul, L., Forsythe, A., Tomiyasu, U., Wechsler, A., et al. (1985). Differentiating characteristics of the human speech cortex: A quantitative golgi study. In D. F. Benson, & E. Zaidel (Eds.), *The dual brain: Hemispheric specialization in humans*. (65-74). New York, NY, US: Guilford Press.

Tucker, D.M. & Williamson, P.A. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review*, 91(2), 185-215.

Tulving, E., Markowitsch, H.J., Kapur, S. & Habib, R. (1994). Novelty encoding networks in the human brain: Positron emission tomography data. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 5(18), 2525-2528.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

Wicha, N.Y., Moreno, E.M. & Kutas, M. (2003a). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483-508.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003b). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

Wlotko, E.W. & Federmeier, K.D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001-3014.

Wolford, G., Miller, M.B. & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *Journal of Neuroscience*, 20(6), RC64.

Yamaguchi, S., Yamagata, S. & Kobayashi, S. (2000). Cerebral asymmetry of the “top-down” allocation of attention to global and local features. *Journal of Neuroscience*, 20(9), 1-5.

CHAPTER 6.
GENERAL CONCLUSIONS

6.1. Goals

The overarching goal of this thesis was to investigate the anticipatory nature of sentence comprehension as a route to better understanding how external language cues tap into semantic organization and trigger meaning construction in the brain. While generative linguists have long argued in favor of more bottom-up comprehension models, a decade's worth of evidence for early, context-based effects suggests that like other neural domains, the brain does not exclusively wait to receive linguistic input to begin processing. Testing for prediction in language comprehension is an interesting question not (only) because we care about the timing of sentence processing, but rather because of what predictive processing may reveal about the probabilistic and experiential nature of information storage and activation, the strategies the parser adopts when sentences continue in more or less expected ways, and the brain mechanisms that may allow for such flexibility. In particular, the research presented in this thesis was designed to address the following questions:

(a) Can we uncover convincing, pre-target evidence for pre-activation of linguistic content during online sentence processing? If so what is the nature of such prediction? At what levels (featural, word-level, semantic, syntactic, phonological, etc.)? Across which degrees of constraint (for highly and/or weakly constraining contexts)? Under what processing limitations (at normal or rapid input rates, or both)?

(b) Are there consequences to mispredicting? Even if preactivation facilitates processing a majority of the time, there are inevitably instances in which committing to a highly likely continuation is not beneficial. Is there evidence of such a cost? How do increasing constraint and the receipt of more or less likely information interact?

(c) Do the brain's two hemispheres utilize contextual constraint in similar ways to preactivate upcoming sentence content? At which stages of processing are potential hemispheric prediction biases evident during online language comprehension?

In the following sections, we will address these questions in terms of our empirical findings and will extrapolate a more comprehensive view of the anticipatory nature of sentence processing. In addition, we will expound our view of some methodological issues relating to experimental design, performing various analyses, and some of the pitfalls of framing research in terms of certain widely-used, but ill-defined, psycholinguistic concepts.

6.2. Upcoming linguistic input is pre-activated to varying degrees via contextual constraint

Though we declared in the Introduction to this thesis that there is mounting evidence for anticipatory language processing, it is a question that has been inherently tricky to address, due to the difficulty in distinguishing between predictive and integrative processes. Online measures (such as ERPs and eye-tracking) afford some means of observing effects prior to receiving target input, but at least for the electrophysiological studies, this still requires some clever experimental design in order to find evidence for events that have not yet occurred. While anticipatory processing has been observed across a variety of neural domains (e.g., somatosensory: Carlsson, Petrovic, Skare, Petersson & Ingvar, 2000; gustatory: Simmons, Martin & Barsalou, 2005; and visual processing: Slotnick, Thompson & Kosslyn, 2005), acceptance of a role for prediction in language comprehension has met with some resistance. Traditionally, bottom-up processing models have dominated the literature, with effects of context allowed for only postlexically, after sensory input has been initiated. Though some have framed arguments about linguistic preactivation in terms of either lexical priming via spreading activation (fast, automatic processing, at short delays, supposedly with benefit but no cost) or conscious processing strategies (slower, at longer delays, with as least as much cost as benefit) – positing that neither type typically guides lexical access during normal, fluent reading – what we argue for in this thesis is something different. We believe we have presented evidence for fast, automatic, pre-lexical, unconscious, probabilistic pre-activation of words (or their features), for which a graded processing cost is evident once the input is inconsistent with that of more preactivated items. We believe that this preactivation can occur on many different levels, though the primary evidence in this thesis is for the preactivation of lexical word forms.

The most compelling evidence for linguistic prediction – constituted by studies which allow disambiguation between integrative and predictive process – has resulted from online comprehension studies utilizing methods such as eye-tracking or ERPS, both of which provide temporally continuous measures. Eye-tracking studies using the visual world paradigm (e.g., Altmann & Kamide, 1999; Kamide, Altmann & Haywood, 2003; Boland, 2005; Knoeferle, Crocker, Scheepers & Pickering, 2005) demonstrated that as function of context, comprehenders will look to the most likely candidate entities present in a scene, prior to receiving the associated input. However, a possible shortcoming of these studies is that potential "predictees" are visually present, which some might suggest makes such findings less generalizable to language processing in the absence of candidate visual referents. On the other hand, in ERP studies such as Experiments 1, 2 and 3B, as well as studies by van Berkum, Brown, Zwitserlood, Kooijman & Hagoort (2005) and Wicha, Moreno & Kutas (2003; 2004), upcoming word candidates were constrained by preceding linguistic context only. In addition to finding evidence for linguistic preactivation at time points temporally preceding sentential targets, Experiment 1 went beyond the other ERP prediction studies in demonstrating that preactivation can be at the level of lexical word forms, and is graded depending on the constraint of the context.

6.3. Pre-activation is probabilistic

We argue that the graded prediction effect at prenominal articles in Experiment 1 suggests that the parser preactivates information probabilistically, i.e., anticipating potential upcoming continuations over wide ranges of constraint. Our findings indicate that preactivation is not limited to only the most constraining contexts, where only a single dominant item is liable to be encountered. Rather, it seems that the processor utilizes whatever degree of contextual constraint is available to predict, even when the pool of likely continuations is greater than one. By design, this might not be considered the most efficient method of processing, since particularly for less constraining contexts, there is a high probability of the upcoming continuation not being preactivated. However, we propose that preactivation is not explicitly (consciously) calculated, but rather comes for “free” as a result of incremental context updating, and as such is not limited to a subset of highly constraining contexts. Inevitably, the predictive mechanism will sometimes “miss the mark”, and indeed we observed an ERP pattern (in the form of the LP) over all the studies in this thesis, which was consistent with predictions being violated – more so for highly constraining contexts, but also when constraint is not as high.

In arguing for probabilistic prediction based on our findings from Experiments 1 and 2, we suggest that preactivation is supported by multiple sources. A longstanding belief within certain psycholinguistic circles has been that brain mechanisms responsible for syntactic structuring during online language comprehension are modular, i.e., encapsulated from other cognitive and perceptual systems. If syntactic structuring is the main (or only) process driving comprehension, it is not so difficult to understand why the concept of preactivation has traditionally seemed untenable: syntactic information alone would not go very far in whittling down the huge number of possible continuations for, say, an upcoming noun. However, as Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy (1995) state:

> More promising are theories by which grammatical constraints are integrated into processing systems that coordinate linguistic and nonlinguistic information as the linguistic input is processed.

The studies herein focus on the role of linguistic information in shaping predictions. We interpret our evidence for probabilistic preactivation as being supported by the functional organization of semantic memory (e.g., knowledge of people, places, things, and events), and propose that preactivations of such information are concurrently shaped by phonological, syntactic, and pragmatic constraints. It is possible that our stored memories built up from everyday experiences, as well as information gleaned from the immediate sensory environment, combine rapidly and dynamically to generate expectancies not just about language, but abut the world in general. Bar (2007) suggests that there is no reason to believe that

> ...predictions at different levels of complexity, abstraction, timescale and purpose use mechanisms that are qualitatively different.

While we limit ourselves to the study of language comprehension in this thesis, a domain-general neural prediction network is nonetheless an interesting proposal to ponder.

Our graded, prenominal article prediction finding in Experiment 1 is in line with a growing list of experimental studies and parsing models (e.g., Jurafsky, 1996; Crocker & Brants, 2000; Hale, 2003; Levy, 2008; Smith & Levy, 2008) that project probabilistic expectations about various aspects of linguistic input during comprehension. The factors that inform such predictions are likely to be numerous and varied, and intersections of experimental testing and computational modeling may offer one promising avenue for determining their individual roles and weights, as well as how they interact with other factors. Determining these contributing factors, as well as investigating the neural mechanisms involved in such predictive language processing, remain open questions and ones ripe for aggressive investigation.

6.4. Pre-activation generalizes across input rates

One argument that has been leveled against ERP language comprehension studies, and one that would seem of particular relevance for studies of predictive language processing, is that using RSVP introduces an artificial delay between words, beyond that of normal reading. It has been suggested that this “extra time” may allow readers to form conscious predictions about upcoming targets. Rayner, Ashby, Pollatsek & Reichle (2004) go so far as to say, in reference to sentence priming studies accompanied by naming or lexical decision tasks, that

> …even if there is not time for conscious prediction…the timing between the intake of information from the prior text and the intake of information from the target word is appreciably different than in normal reading.

One can imagine that the same argument might be used against our studies, though one of the primary benefits of ERPs is that at least in terms of not requiring an additional task, they generally offer a more naturalistic ways of examining linguistic phenomena. Implicit in the above statement is the idea – which we acknowledge to be true – that it is the timing of evidence for prediction which is most crucial. After all, no one disputes that information can be activated based on contextual cues (i.e., in a cloze norming task, where no time limits imposed, individuals converge to varying degrees on appropriate sentence continuations); however, the crucial question seems to be one of a) *when* does prediction occur, and b) *what* is meant by prediction. On the second point, we have attempted to be clear throughout this thesis about what it is that we mean by prediction (aka preactivation or anticipation, as evidenced at some sufficiently prelexical time point), and hope that it has been apparent why cloze norming is *not* representative of the kinds of processing we would consider as preactivation. On the first point, although we outlined in Experiment 1 our explanation for why our finding of a prenominal prediction effect does address the "when" question, we nonetheless probed this issue more directly in follow-up Experiment 2.

While our prediction-related findings in Experiment 2 were non-identical to those observed in the original study, we nonetheless observed effects of prediction at the prenominal articles. These effects were most pronounced at the extreme ends of the target nouns' cloze range, although a subset of participants – those whose offline testing scores indicated they may be more experienced or facile readers – exhibited graded prediction. We offered two possible explanations for the difference between prediction effects at the two presentation rates. The first is that the prenominal articles are likely to have different information value in high versus less constraining contexts, more confirmatory of upcoming information in the first case and more critical for shaping predictions in the second. We suggest that under speeded processing, the ERP at the article may be sensitive to only the more well-formed noun predictions in sentences with high constraint. The binary results we observed are consistent with previous prediction-related research using similar experimental paradigms (Wicha et al., 2003 and 2004; van Berkum et al., 2005). We also proposed that an alternate explanation for the differences in results at the two presentation rates had to do with the problem of overlapping ERP components at the 3.3 words/sec rate; in particular, the N400 time window of the preceding article (200-500 ms post-article onset) overlaps with the first 200 ms of the subsequent noun. Hence, there was potential for early, attention-related effects at the noun to dampen the article prediction ERP effect. van Berkum et al. (2005) avoided this issue first, by presenting stimuli in the auditory modality, and second by separating the prenominal gender-marked words from their noun targets by an intermediary phrase. Eventually, we, too, plan to run an auditory version our experiment, though coarticulation effects for our particular *a/an* paradigm could potentially lead to a different set of analysis difficulties.

Ultimately, what we would like to be able to demonstrate is that preactivation occurs incrementally, with different factors exerting their influence over different time scales, at different points in a sentence. It is easy to imagine that a particular word may be pre-activated most strongly immediately before it is about to be read or heard, after a large amount of context has already accrued. Indeed, in Experiments 1 and 2, evidence of pre-activation was detected (by design) at the word immediately preceding the target noun. A more informative finding, however, and one ultimately more supportive of the argument that linguistic information gets pre-activated as a natural consequence of language processing, would be to show that predictions are formed incrementally and probabilistically, well before the confirmatory input is received. What is needed is a way of showing that pre-activation of a word can be altered over the course of a context – in other words, to find evidence for modulation in preactivation levels as a context unfolds and as constraint strengthens. Admittedly, this is no easy task; however, such a finding would certainly lend support to the idea that preactivation occurs unconsciously, probabilistically, and over time.

In sum, our findings of prediction effects across a range of normal input rates – in combination with similar evidence from auditory modality prediction studies – support the proposal that such effects are not likely due to post-lexical (conscious) processing. Against psycholinguistic theories maintaining that during online comprehension contextual information has its effect only at (during or after, but not before) the processing of a given lexical item, our N400-like effects observed at the prenominal articles in both Experiments 1 and 2 were evident as early as 200 ms post-article onset – either 300 ms or 100 ms (respectively) prior to the target noun appearing. For those who might argue that such results are an artifact of the spare time afforded by RSVP, it is not clear what the necessary timing/presentation rate cutoff would be in order for prediction to be acknowledged. We believe that our findings are sufficient for arguing a case for preactivation, though further investigation of the timing-related differences between Experiments 1 and 2 will undoubtedly inform these results.

6.5. There is an apparent processing cost that increases with constraint violation

With Experiments 1 and 2, we began our investigations of anticipatory language comprehension by setting up an experimental paradigm which allowed us to observe prediction effects prior to more or less constrained sentential targets (nouns). If ERP differences were observed to semantically equivalent prenominal indefinite articles (*a/an*), then this would provide strong evidence that the nominal word forms had already been preactivated. Our results supported this hypothesis. In addition, in these studies, as well as Experiments 3B and 4 of this thesis – all of which used similar types of sentence stimuli with target nouns preceded by *a* or *an* – there was another consistent, though surprising, ERP finding. This effect took the form of a late positivity (LP) following the target noun N400, more anterior than posterior, and with a sometimes more leftish hemispheric bias. In Experiments 1 and 2, we observed that this LP was more pronounced to low cloze than high cloze noun continuations, and exhibited a graded increase in amplitude as noun cloze decreased over certain scalp areas. These LP/noun cloze correlations were present at both rates of presentation but were stronger at the shorter SOA. Post-N400 late positivities are not unheard of, however the more frontal nature of our LP effect in combination with its graded relationship with decreasing cloze probability, was something novel. Our first attempt (Experiments 3A and B) at systematically probing this fortuitous finding was to test how a particular aspect of expectancy might be getting violated; specifically, we wondered if when prediction-inconsistent prenominal articles were received, whether comprehenders' expectancies were revised to include a phonologically consistent adjective in order to salvage the expected noun. We tested this idea by contrasting different low cloze noun continuations preceded by either prediction-consistent or inconsistent articles, with the hypothesis that we might only observe the LP to the low cloze nouns preceded by the prediction-inconsistent articles, because only in these cases would there be any cause to insert an adjective. We, in fact, did observe an increased LP to the low cloze relative to high cloze nouns, however, contrary to our hypothesis, this effect was present for *both* types of low cloze noun conditions. This led us to surmise that the LP effect was not related to a violation of syntactic expectation, per se. Within the same study, we probed further by opportunistically exploiting a range of contextual constraint within our sentence stimuli to investigate whether the LP effect might instead be related to constraint (expectancy) violation, in a broader sense. Our revised hypothesis was that if the LP was indexing some functional process related to highly preactivated information being disconfirmed, then we should observe larger LPs to low cloze nouns continuing highly constraining contexts than to low cloze nouns continuing less constraining ones. Indeed, this is the pattern we observed, over alternate methods of quantifying constraint violation. Because cloze probability and constraint are confounded over their upper ranges, holding cloze probability constant while manipulating constraint level allowed us dissociate the influence of the two factors. Our results also indicated that similar to the prediction effect present at the prenominal articles, the LP constraint violation effect was also graded.

Throughout this thesis, we have remained intentionally agnostic about the functional nature of the observed LP effects. We have gone only so far as to suggest that the effect appears to be related to violation of constraint, though we have not specified whether it might be reflecting aspects, for instance, of inhibition, reanalysis, revision, conflict monitoring, etc. Adjudicating such possibilities is beyond the scope of the current set of studies. We will, however, examine the implications of interpreting these results as a kind of "misprediction cost." In recent years, there has been a wave of studies in which late ERP positivities of various time windows, durations, scalp distributions, and eliciting conditions have been observed to so-called "semantic" manipulations. These findings of "semantic P600s" have been somewhat surprising, and have brought into question not just the traditional division between supposedly semantic and syntactic processes (reflected by N400 and P600 ERP components, respectively), but also the contrast between how experimenters formulate and define their experimental conditions, and the (potentially fuzzier) lines along which the brain may distinguish these phenomena. To suggest that all P600s (aka late positivities) observed across both the traditional syntactic as well as more recent semantic manipulations are reflecting exactly the same kind of processing, or even that they are members of the same family of components, may be overstating things, especially without many direct, within-study comparisons of effects for the different types of eliciting conditions. However, we believe our "misprediction cost" explanation of the LP is more compatible with theories suggesting that such effects are best accounted for in terms of cognitive control and conflict monitoring. Kolk and colleagues (Kolk, Chwilla, van Herten & Oor, 2003; van Herten, Chwilla & Kolk, 2006; Vissers, Kolk, van de Meerendonk & Chwilla, 2008), for instance, have suggested that when a conflict appears between incompatible sentential representations, reanalysis is initiated to check whether the conflict is due to processing error. Novick, Trueswell & Thompson-Schill (2005) and Thompson-Schill (2005) suggest that such reanalysis might be due to selecting between competing representations based on task demands. Though none of these studies frames an LP hypotheses specifically in terms of preactivation or prediction violation, we believe our results, and others, may be compatible with this proposal. Our suggestion is that the “conflicts” that arise need not be ones of syntactic or even semantic violation; rather, the conflict can arise when items or their features are preactivated to varying degrees, but disconfirmed by the physical input. This more linguistically domain general proposal for the LPs is also consistent with the view that the cognitive control mechanisms involved in sentence comprehension may be similar to those employed in more general conflict tasks like the Stroop task (e.g., Novick et al., 2005; Thompson-Schill, 2005; Ye & Zhou, 2008). In addition, our observation of a generally more frontal scalp distribution to the LP effect is roughly consistent with imaging data that have suggested possible roles for various frontal and prefrontal cortical areas in inhibition (e.g., Aron, Robbins & Poldrack, 2004), error detection (e.g., Rubia, Smith, Brammer & Taylor, 2003), and suppression of interfering memories (e.g., Anderson, Ochsner, Kuhl, Cooper, Robertson, Gabrieli, et al., 2004). While ERP patterns at the scalp do not allow for direct mappings to specific brain areas, the frontal nature of our LP, in conjunction with these localized neural functions, is nonetheless suggestive.

In terms of language processing theories, it has been widely argued that because sentence contexts in natural language are rarely sufficiently constraining (Morris, 2006; Gough, Alford, Wilcox, 1981; Mitchell, 1982), anticipatory processing of upcoming words would seldom “get it right” and that “getting it wrong” should ultimately be costly. In fact, a paucity of evidence for processing cost has likely been a primary source for resistance to anticipatory language models. We would argue, however, that even a lack of evidence for a processing cost could be consistent with an anticipatory model in which subtle, incremental, unconscious, multilevel predictions are continually recomputed with each new bit of (multi-sourced) input, and for which this process occurs without penalty. Our data, however, suggest that at least under some circumstances, such costs may be evident when predictions are not met. This cost need not reflect inhibition or revision - though it might – but instead may represent an allocation of resources to "shift frames" in order to bolster an already weakly activated representation or activate a new contextual representation. Undoubtedly, this proposal requires further investigation.

As a final note, we should point out that although we have framed our LP finding in terms of reflecting a potential "cost" for misprediction, this may not be the only – nor possibly the correct – interpretation. There are, however, various reasons why we did so. The first is that the idea of "cost" in sentence comprehension, as for other neural operations, implies more effortful processing leading to a decrement in some behavioral performance measure. For instance, behavioral language comprehension studies have historically mapped the idea of "cost" to findings of temporal delays (e.g., in lexical decision and word naming tasks), reduced accuracy in comprehension questions, and/or decreased memory recognition or recall performance. With regard to theories of linguistic prediction, the lack of such behavioral evidence in early off-line sentence processing studies where constraint violations should have led to such decrements, was indeed taken as a strong argument against anticipatory models (e.g., Fischler & Bloom, 1979). There is a long tradition, then, carried over from the behavioral literature, of thinking about post-stimulus effects to constraint violations in such terms. Even with an on-line methodology such as eye-tracking, increased regressive eye movements and longer fixation times in reading studies are relatively interpretable in terms of "costs" to encountering items that are unanticipated based on their preceding contexts. With ERPs, however, it is certainly more of an interpretative leap to link modulation of a component's amplitude to any type of decrement in behavioral performance, because none has explicitly been demonstrated. A second motivation, however, for considering our LP effect as reflecting some type of misprediction "cost" stems from its directional and temporal similarity to the P600, a component also generated to linguistic stimuli, though most typically to more syntactic manipulations. Indeed, amplitude variation in the P600 has been linked with the "cost of reprocessing" (Osterhout, Holcomb & Swinney, 1994), and for this reason we were inclined to extend the general idea of "cost" to our more semantic manipulations. Clearly, our effects were not syntax-related, but in light of the recent wave of studies challenging the P600's specificity to syntactic processing, the proposal seemed a reasonable one. Ultimately, there could be other functional interpretations of the LP that may not involve any sort of "cost". For instance, a consequence of receiving surprising information could be that some sort of learning mechanism has to update or recalibrate the predictive value of various inputs or their combinations. Or it could be that the LP reflects the brain's graded sensitivity to conflicts between expected and encountered information, but that there is no "trade-off" in the (presumably additional) cognitive load imposed by such inputs. Adjudicating between the possible explanations will undoubtedly require clever experimentation, and will likely be informed by more general theories of neural prediction.

6.6. There is a double dissociation between the sensitivities of the N400 and the LP: opposite sides of the same coin?

Across the four ERP studies described in this thesis we found evidence for a double dissociation between the functional processing reflected in the N400 and in the LP ERP patterns. As has been shown repeatedly across a multitude of paradigms, N400 amplitude decreases in a graded fashion as a function of how contextually probable a continuation is considered to be – a value which has primarily been determined in terms of cloze probability ratings. The component is not an anomaly detector, per se, though it is sensitive to such information. Rather, the N400 seems to reflect the benefit of contextual facilitation, from linguistic as well as nonlinguistic sources. The N400 has not, however, been show to differentiate between unlikely sentential continuations to highly and weakly constraining contexts; thus, regardless of “violation” level, N400 amplitude remains constant. This observation, (as mentioned above) has led some to posit that preactivation is not the mechanism by which normal language comprehension occurs.

On the other hand, the LP effect observed in our ERP studies, which upon initial analysis seemed to exhibit a sensitivity to cloze (increasing in amplitude with decreasing cloze) was revealed to differentiate between precisely the cases that the N400 did not; namely, low cloze continuations to high and low constraint contexts. Indeed, these are the cases in which cloze and constraint violation dissociate. So while N400 amplitude decreases reflect the benefit of contextual support, LP amplitude increases appear to reflect a cost to preactivating but not encountering such information.

One interesting finding over our studies was that what we have referred to as the late positivity, may actually be initiated within the same time window as the N400 (prior to 500 ms). The early onset of this effect was evident particularly over left anterior scalp sites (where LP effects in the later time window tended to be largest), while the posterior scalp areas continued to exhibit more N400-like patterns throughout the typical N400 time window (200-500 ms). This temporal overlap may offer one potential explanation for why LPs have not been consistently observed across traditional N400 studies that use paradigms similar to ours (though there are numerous other reasons why this might be the case). Indeed, the overlap of the two ERPs may be more extensive than we are able to observe, because it may be that LP effects are masked by generally stronger posterior N400 effects. These findings also suggest that the processing associated with the N400 and LP may not be serial, with LP-related processing capable of being initiated without N400-related processing having concluded. If LP processing is reflecting some type of reanalysis, then the temporal overlap with the N400 might suggest that an item need not be fully integrated in order for this process to begin.

6.7. Both cerebral hemispheres are sensitive to message-level constraint

One of the goals of Experiment 4 was to compare whether the brain's two hemispheres benefit in a similar manner from gradations of message-level contextual constraint. Results across previous studies and methodologies have ranged from arguing that only the LH is sensitive to sentential constraint, to proposing that both the LH and RH build up contextual representations, to suggesting that the cooperation of the two hemispheres may be required to reveal the graded N400/cloze probability correlations that have been observed across central presentation studies. Our study indicated that both hemispheres individually exhibit sensitivity to varied probability sentence continuations over different levels of constraint. Both LVF and RVF presentation led to similar sentential N400/cloze probability correlation patterns, of comparable strength and distribution – patterns which paralleled those observed for central presentation. Though our N400 findings at the noun cannot adjudicate between integrative or predictive comprehension models, the outcome is nonetheless a useful one to determine, particularly in light of interpreting the prediction-related ERP effects that may occur before and after processing of sentential targets in the paradigm used here. In Experiment 4 we did not test the prenominal ERP prediction effect by lateralizing the prenominal articles, because we first wanted to establish the nature of effects related to processing the target nouns, which we believed had more of a precedent and would be larger than potential effects at the article. We also considered determining the noun effects a prerequisite because lateralized constraint/cloze probability effects had never been examined over such finely graded ranges of these factors. Our findings indicated that though (as will be discussed in the following sections) modulations of constraint and cloze were processed non-identically by the two hemispheres over a later time window, processing of sentential noun targets during the N400 time window was generally similar, which is the outcome we had predicted. These lateralized effects were consistent with those of Kutas & Hillyard's 1984 (central presentation) study, which first established the inverse relationship of cloze and N400 amplitude in central presentation, and for which the component was determined not to index constraint violation. Thus, sentential semantic integration appears to benefit from facilitative context for both cerebral hemispheres.

Our Experiment 4 findings of similar N400 amplitude/noun cloze correlations for presentation to both VFs stand in opposition to theories such as Beeman's coarse coding hypothesis (Beeman, 1998), which suggest that semantic activations are broader and weaker for RH than LH processing. The Experiment 4 noun N400 results indicated that both hemispheres were able to take advantage of fine-grained, message level constraint, which accrued as a result of build-up of contextual representations, rather than through lexical-level priming – evidence that also refutes theories like those of Faust and colleagues, who have suggest this is the primary processing strategy upon which the RH relies. Another possible outcome for the noun N400 findings could have been that one hemisphere, but not the other, showed a graded sensitivity to cloze probability, which would have suggested that the centralized pattern was obscuring processing differences between the two hemispheres, as Wlotko & Federmeier (2007) observed. However, this pattern was not evident in our results, which were elicited from stimuli with a wide range of cloze values, which were sentence medial, which were all congruent, and for which the ratio of low cloze to medium and high cloze continuations was relatively low (around a third of the items). It is unclear whether any of these factors alone or in combination may have contributed to differences in the two studies, though Wlotko & Federmeier themselves suggested that different results might be obtained using stimuli with a more continuous range of cloze probabilities.

6.8. LH biased sensitivity to constraint violation supports a model of a more anticipatory LH sentence processing

Building on one of the more interesting findings in this series of studies – that of increasing LP to unexpected continuations in constraining contexts – in Experiment 4 we observed that the LH was more sensitive than the RH to such violations of sentential constraint. This effect, while graded with level of constraint violation, was strongest when highly constraining contexts were followed by low cloze continuations, e.g., '*For the snowman's eyes the kids used two pieces of coal. For his nose they used a berry from the fridge*', where *carrot* is expected. The LH bias for constraint violation processing is in line with proposals by Federmeier & Kutas (1999) and Federmeier (2007) that LH comprehension involves more top-down contextual processing, and functions more predictively by preactivating information based on associations in long term semantic memory, prior to encountering information.

It has been an open question, with regard to anticipatory comprehension models such as Federmeier's and others, as to whether there is a consequence to not encountering anticipated input. For instance, it could simply be the case that the prediction landscape is modified – without cost – with each new bit of input. Or it could be that constraint violations are observable in the ERP waveform only at points in a sentence where it is clear that the semantic course of a context has been altered beyond needing to keep such information active in working memory. For example, in a stimulus sentence (glossed here from the original Dutch) like one used by van Berkum et al. (2005), '*The burglar had no trouble locating the secret family safe. Of course, it was situated behind a big but unobtrusive painting*', when comprehenders encounter the determiner *a*, preactivation of the word *painting* may increase dramatically. When the next word '*big*' is encountered instead, this information serves to both deny the immediate prediction for the upcoming word as well as signaling that the next word encountered is now very likely to be *painting*. As the next word *but* comes in, *painting* is again disconfirmed, although the following word *but* may again serve to deny local prediction while continuing to feed into the expectation that *painting* will eventually be received further downstream. This example is meant to highlight the idea that the kind of preactivation we are proposing is a mechanism that works incrementally and probabilistically, and for which the items that have received the greatest preactivation are not necessarily the ones confirmed by the immediate input. In this sense, a "cost" function might not seem to make much sense. However, one can imagine that preactivation "costs" might not occur with every new piece of linguistic input, because reanalysis, revision, inhibition, or conflict monitoring (whatever the LP-inducing function might be!) may only be necessarily when the overall preactivated contextual "trajectory" is violated. *Big* and *but* may not induce "costs" because it is still likely that *painting* is on the sentential horizon.

As our data from Experiment 4 indicated, the LP effect which we have suggested relates to constraint violation was more striking for LH processing, though not nonexistent for the RH. While the LH showed a more graded sensitivity to constraint violation, the RH effects were evident only for highly constraining contexts followed by low cloze continuations. However, the RH LP effect was completely eliminated when constraint violation was calculated in a manner which took into account how the expectation for a particular noun was affected by receipt of a prediction-inconsistent article – information which served to alter the more local prediction. While the LH seemed more likely to "ignore" this information to the extent that a high constraint level might still be maintained in the face of an unexpected article, the RH seemed to adopt a more veridical approach, appearing to downgrade constraint with receipt of a low cloze *a* or *an*. We have suggested that this pattern of results may be consistent with the view of the LH as an "interpreter" (i.e., responding more to the perceived probability of an event occurring even if the event's occurrence is not immediate), while the RH is less likely to generalize away from the input (Gazzaniga, 2000). These factors may well shape the manner in which preactivation occurs in the brain's two hemispheres, though undoubtedly further research is required.

6.9. Graded constraint violation effects are consistent with graded prediction

Contrasting high versus low cloze sentence continuations in the LP time window, for both LH and RH processing we observed the largest constraint violation effects for highly constraining contexts. However, when constraint violation effects were calculated over the full range of constraint, there was a clear gradation in LP amplitude to low cloze words as a function of constraint. The correlational nature of this effect is relevant for two reasons. The first is that it signals, similar to our prediction findings, that constraint violation is not all or none. So even when the possibilities for upcoming sentential input have not been narrowed to a single highly likely continuation, there still appears to be some kind of processing cost for encountering input which has not been preactivated by preceding context. These graded effects suggest that prediction and constraint violation-related processing are systematic, reflecting the default way that the brain processes a large range of stimuli, rather than being a more "strategic" approach that kicks in only under certain (e.g., highly constraining) circumstances.

The second point – which is directly related to the first – is that if preactivation is indeed probabilistic and itself a graded process, then a constraint violation cost that is also graded is ultimately highly compatible. Though it seems likely that the brain regions involved in generating linguistic predictions are different areas than those responsible for the LP violation effect, they are nonetheless tandem processes. The larger and more consistent graded prediction violation effects in the LH accord with models of this hemisphere engaging in more anticipatory sentence comprehension. A planned follow-up study in which we use our same lateralized experimental paradigm, except this time presenting the prenominal articles to the VFs, could potentially provide results that would round out the LH dominant prediction/constraint violation model.

6.10. Graded experimental factors and analyses offer insights beyond those afforded by binary ones

Some of the most novel findings in this thesis emerged through the use of correlation analyses. From our finding of a graded N400 prediction effect at prenominal articles in Experiment 1, to correlations of LP amplitude with graded constraint violation in Experiments 3 and 4, our findings went beyond the kind of binary comparison at extreme ends of a particular experimental factor typically performed in psycholinguistic experiments. In particular, the use of stimulus sets that took advantage of graded ranges of prediction-related cloze probability or constraint violation allowed us not only to investigate the sensitivity of the N400 and LP to such factors, but allowed us to strengthen our arguments that the processing reflected at the upper range of cloze (or constraint violation) was the same kind of processing – just to a variable degree – as that across middle and lower ranges. Graded effects of cloze probability and constraint violation also indicated that associated processing was not of the sort in which some threshold must be reached to achieve an effect.

Importantly, for Experiment 1, our finding of graded N400s at prenominal articles argues strongly for the probabilistic preactivation of linguistic information. At the same time, our correlation data allow us to rule out a form of prediction in which across contexts only a discrete item gets preactivated; in other words, anticipation does not occur just in highly constraining contexts, but also in sentences where there may be no single, dominant continuation. These results are more in line with connectionist models that argue for a landscape of prediction in which the parser is continually updating and estimating the likelihood of upcoming material (e.g., Elman 1990, 1995; Seidenberg & MacDonald, 1999).

Our confidence in arguing for similar N400-type effects at both the article and noun in Study 1 was also strengthened by the comparable correlation values and scalp distribution patterns over both word types. Due to the inverse problem (any single distributional pattern of ERP effects at the scalp can be attributed to multitude of neural generators), it is always challenging to argue that two ERP patterns are reflecting similar cognitive processing. The cloze-graded nature of both the article and noun ERP effects (in addition to the timing and distributional similarities) make our interpretation less disputable.

In addition to the article prediction effects, we observed graded N400 mean amplitude at more and less expected nouns in all the studies in this thesis. While for centralized presentation (at RSVP rates of both 2 words and 3.3 words/second) this was not a novel finding, the high noun cloze/N400 amplitude correlations nonetheless confirmed that the more and less probable content words were being processed in the expected way. Our results for lateralized presentation were less precedented, as the hemispheric inverse relation of cloze probability and N400 mean amplitude had only previously been revealed for more discrete analyses (e.g., using either two or three levels of cloze probability). Our finding of similar correlation patterns for nouns presented to both the left and right visual fields indicates that both cerebral hemispheres are sensitive to cloze probability, and that the N400 pattern observed in centralized studies is not necessarily an amalgam of the two hemispheres' contributions (as proposed by Wlotko & Federmeier, 2007).

Notably, in addition to the N400 correlation patterns over all the ERP studies presented herein, we noted strong correlation values in various post-N400 time windows, which we referred to as the late positivity (LP) time windows. In both Experiments 1 and 2, we correlated LP mean amplitude with cloze probability and noted that with decreasing cloze, LP mean amplitude tended to increase. More systematic manipulation and testing of this LP effect in Experiments 3 and 4 revealed that the component's amplitude increased not specifically as a function of cloze probability, but rather with the degree to which a noun continuation violated the constraint of a particular context. In other words, low cloze continuations to high constraint contexts led to larger amplitude LPs than low cloze continuations to medium and lower constraint contexts. This sensitivity was evident when constraint violation was calculated in two different ways. Again, the graded nature of this effect led us to surmise that similar to the graded effects of prediction noted at the prenominal articles, our LP effect may be reflecting the degree of violation of such probabilistic, contextually-based pre-activated material. The results of these LP correlation analyses also offer a powerful point of comparison for further exploration of the functional nature of this component, to the extent that, for instance, we could test whether our LP is related to other so-called "semantic P600" effects observed in the literature by exploring whether graded manipulations of the experimental factors on which those results are suggested to be based (e.g. subcategorization biases, thematic role and animacy violations) obtain.

While the kind of topographically distributed scalp correlation maps that we have presented in this thesis have allowed us to draw some powerful conclusions about the cognitive ERP effects we have observed, we acknowledge that these examinations would benefit from more statistically rigorous methods, for instance, in comparing correlation strength and distribution patterns across, e.g., different time windows, different participant groups, etc. The principal component analysis regression approach employed for the comparison of lateralized N400 effects in Experiment 4 offers one approach. In general, though, our hope is that more psycholinguistic studies will begin incorporating graded experimental manipulations and methods of continuous analysis, and that standard methods of comparison will emerge.

6.11. Beyond cloze probability and constraint

In language comprehension research, it is not uncommon for researchers to refer to context, constraint and cloze probability all within the same study. So in which ways do these concepts overlap, or better yet, how are they differentiated? First, in the sense used for many language experiments, a language context can range from being highly constraining (maximum constraint would mean that a context would have a single, plausible continuation) to weakly constraining (a context could plausibly continue in a potentially unlimited number of different ways). In the case of a high constraint context, the constraint value (e.g., 95% constraint) is most frequently determined by (and thus has the same value as) the highest cloze item (95% cloze). The same process determines the constraint value of a low constraint context (e.g., a context for which the most common continuation is provided by only 12% of respondents also is said to have 12% constraint). So when researchers conduct studies in which they create sentence stimuli utilizing the highest cloze continuations for each context, they must keep in mind that the factors of constraint and cloze probability are confounded. Why is this relevant? Well, in many instances it is not – unless one is interested, for instance, in issues of prediction versus integration, or if a researcher is concerned with investigating what happens when a sentence does not continue in an expected way. This is because constraint and cloze probability are only dissociable for sentences which do not continue with the highest cloze item for that context (e.g., a context with 80% constraint continues with a word with only 5% cloze). In examining prediction and constraint violations within this thesis, we have compared low cloze items in high constraint (e.g., *He mailed the letter without a…thought*) to those in low constraint (e.g., *He was interrupted by a…thought*) contexts. With low cloze probability held constant, one can more clearly examine the contribution of the factors leading up to, as well as factors triggered upon presentation of, the target item.

We believe that one point highlighted by the studies included herein is that widely used terms like “cloze” and “constraint” are worth breaking down. While these concepts are generally taken at face-value, their use may mask potential sources of interesting variability which could ultimately be problematic when attributing particular psycholinguistic phenomena to their modulations.

6.11.1. Cloze probability

In the study of semantic sentence comprehension, teasing apart the influence of “cloze probability” (the percentage of individuals completing a particular context with a particular continuation: Taylor, 1953) and sentential constraint (the degree to which a context narrows the fan of likely continuations) is not a completely straightforward enterprise. For starters, one must keep the terminology straight. For instance, when addressing research questions relating to language comprehension, concepts sometimes get muddled when “cloze probability” and “expectancy” are equated. Cloze is an offline measure that serves as a proxy for predictability of an item during online processing of a context. What we suggest prediction implies is some type of contextual activation of an item(s) or its/their features prior to encountering it (them) or being forced to produce a continuation. Thus, there is a mapping involved in describing sentence continuations as “expected” or “unexpected” in terms of their cloze, particularly when the research question of interest focuses on the timing of semantic activation in sentence processing, i.e., prediction. Labeling items as "expected" based on cloze norming need not be equated with the item being preactivated (due to the integration/prediction conundrum argument outlined earlier in this thesis). Additionally, cloze values derived from norming are essentially production measures, although they are often used to control language comprehension experiments. Even though it is generally assumed that the two are highly correlated, they need not be identical.

Though in practice we presume to be tapping similar processing via our off- and on-line measures, we cannot be certain that this is the case – particularly due to the experimental conditions under which norming data is typically collected. For instance, in a cloze norming task, there is usually no time limit for participants to supply continuations for truncated sentences, and consequently an item consistently produced across participants may be labeled "expected" regardless of whether that item was produced by individuals after one second, one minute, or – hopefully not – one hour. However, outside of self-paced reading studies, rate of input in comprehension experiments – as in real life – is mostly out of a listener or reader's control. Nonetheless, there is strong evidence of a tight relationship between cloze and measures thought to be related to online expectancy, e.g., N400 mean amplitude (Kutas & Hillyard, 1984). Note, though, that even as early as Kutas & Hillyard's study, it has been known that other factors, such as the strength of an item's featural or semantic relation to an expected item, have the potential to reduce N400 amplitude to otherwise contextually implausible – i.e., low cloze – items.

A point that emerges from the studies described in this thesis, then, is that traditional measures of cloze probability have the potential to confound factors that could ultimately account for important sources of variability in the dependent measure. Because cloze norming procedures typically do not impose any time limits on producing sentence continuations, nor are reaction times or difficulty ratings typically recorded during testing, it is not so hard to imagine that at both the upper and lower ends of the cloze scale (ranging from 0-100%), one could observe very different types of responses that result in similar cloze probability ratings. Consider, for example, sentences (1)-(4) below:

High constraint contexts/high cloze probability continuations:

(1) *He mailed the letter without a...**stamp***

(2) *The fourth letter of the alphabet is...**"D"***

(3) *The IBM logo is often identified with the color...**blue***

(4) (Arguably high constraint): *Violet began cooking her omelet by cracking a...[large, chicken, raw, handful of, few]* ***egg(s)***

Though as a result of offline testing, norming participants may settle on a single high cloze continuation for each of these contexts, it is not clear that listeners or readers are engaging in equivalent processing at targets like these during online comprehension. While sentence (1) is more of a "classical" example of a high constraint context whose continuation reflects the "first word that comes to mind", for the other examples one can imagine that in a norming task, responses – though they may ultimately converge across participants – may require more conscious processing and more time to produce. In sentence (2), this may be because the information has to be accessed serially, in sentence (3) it could be that the kind of world knowledge needed to continue the sentence is not immediately available and requires more extensive memory search. In sentence (4), although the context is highly constraining up until the very end, respondents may be stymied by the indefinite article, and though *egg* will still end up being a high cloze continuation (if one chooses to take intermediate words into account when calculating cloze), the prenominal modifiers used to salvage the noun may be quite varied.

On the flip side, there is also enormous variability between the kinds of sentence continuations which are sometimes included under the umbrella of being "low cloze". Consider the following context and continuations:

(5) *The cop shot his assailant with a...*

high cloze: (a) *gun* (expected), 95% cloze

low cloze: (b) *pistol* (semantically/event structure related), <1% cloze

(c) *cannon* (categorically related), < 1% cloze

(d) *blender* (unrelated/anomalous), < 1% cloze

(e) *glance* (unrelated/frame shift), <1% cloze

While a careful experimentalist might suggest that this potential confound need not be such a problem if various factors are controlled for within a particular experiment, issues arise at the level of between experiment comparisons, or when theories are structured around various interpretations of "cloze" results. The use of standard cloze probability ratings in the psycholinguistic literature is ubiquitous, but rarely are these measures examined in any systematic way. And with good reason – namely, there traditionally has not been a viable alternative to quantifying an item's contextual probability. (Recent work by Roger Levy and colleagues – e.g., Levy, 2008 – suggests that the calculation of an item's "surprisal" value may offer a different way of quantifying conditional probability within a context. Unlike cloze, surprisal is on a log scale, which for one, implies that similar absolute differences in predictability should have a greater impact on difficulty at the lower rather than higher end of the scale. The theory also projects predictability differences even for those sentence continuations that are not the most likely ones within a given context.) In addition, variability between low cloze continuations is difficult to quantify in part because of the limited participant size of typical cloze norming. Eliciting continuations from an order-of-magnitude larger number of subjects would undoubtedly reveal more subtle differences at the low end of the cloze scale, but in general this is not feasible or practical. Some studies, like those of Kutas, Lindamood & Hillyard, 1984 (‘*The pizza was too hot to eat/drink/cry*’) and Federmeier & Kutas, 1999 (‘*He caught the pass and scored another touchdown. There was nothing he enjoyed more than a good game of football/baseball/monopoly*’), have probed semantic relatedness of context anomalous continuations to expected endings, contrasting these types of “low cloze” endings with unrelated, anomalous endings. These studies revealed that cloze alone does not predict of N400 amplitude. Probing different facets of relatedness of low cloze endings to expected ones and different “flavors” of low cloze endings with each other seems like a productive way to examine both the organization of semantic memory as well investigating the types of information that may be preactivated by a given context.

Another issue regarding quantifying cloze probability is exemplified by responses given to a context such as sentence (4) above. While norming participants universally supplied “egg” as a head noun, the preceding modifiers are quite varied. In this case, cloze might be calculated in several different ways (e.g., using the first word only, using an entire noun phrase, counting any answer which contains the head noun anywhere in the response, etc.), which is in fact how we attempted to deal with such results from our own norming data in the studies included in this thesis. An approach that is sensitive to these different levels of similarity and difference inevitably leads to a more thorough understanding of the evoked brain responses and mechanisms of comprehension. As demonstrated in Experiment 3B, this nontrivial difference in calculating contextual constraint based on an alternative quantification of cloze led to our observation of LP amplitude correlating with constraint violation. We believe cloze norming information (across various collection methods and different types and depths of analysis) to be a rich data source – one that could be used to more thoroughly probe associated online measures as well as being informative as a stand-alone measure or to motivate future experimental testing.

Finally, it may be that in studies of prediction in language processing, thinking has been limited by our reliance on the notion of a mental lexicon filled with dictionary-type word entries and by our reliance on the offline cloze norming as the only index of expectancy. It is unclear how similar the process of generating continuations based on prior context is to how we process words online, during normal comprehension. Cloze norming responses may very well prove to be completely relevant to receptive language processes – indeed, ERP evidence assures us that there is some link between the two. However, cloze is unlikely to be the full story, most certainly underestimating prediction across a variety of featural levels, for instance.

6.11.2. Contextual constraint

Though we have detailed above some pitfalls of using cloze probability measures as a proxy for online predictability as well as potential differences in how the measure could be calculated, a perhaps even greater concern centers on generic reference to "contextual constraint". Ultimately, constraint could reflect the contributions of many sources (e.g., syntax, word frequency, concreteness, phonology, world knowledge, etc.), with all of these factors potentially leading to a particular word being more or less expected. Generally, though, in sentence processing studies the influence of individual factors is not specified. For instance, in our own norming data (Experiment 3A) we observed that for a truncated sentence such as, "*Because they were playing baseball so close to the house, the children ended up shattering an…*", the presentation of the indefinite article *an* drastically altered the "constraint" value of the context from when normed with the expected article *a*. It is clear, then, that in an ideal world, "constraint" would be considered on a variety of different levels, with, for example, the cloze value for a particular part of speech, a specific phonological pattern, or a set of shared semantic features being calculated separately from norming results. In addition to these concerns, there is also the issue of how "constraint" (operationalized through cloze norming) underestimates the probability that certain items have a higher likelihood of appearing than their cloze values might reflect. This is particularly true of low cloze items, which though they may not be supplied when testing a limited body of norming respondents, might occur in natural language with a level of frequency higher than, say, implausible continuations.

Quantification of constraint based on cloze values also has the potential to confound contexts that exhibit a small fan of norming responses with those with larger fans. For instance, calculated in the typical way, a context could exhibit 50% constraint either because there are two relatively probable continuations, each with 50% cloze, or because there is one more probable continuation with 50% cloze and a variety of other continuations with lower cloze. There is something inherently different about these two cases, though both would be classified similarly and would unlikely be controlled for in most experimental designs.

As for the "context" part of "contextual constraint", in addition to referring to the physically supplied stimuli themselves (including, but not limited to, written, spoken, signed, gestured, or depicted items or combination of items), within studies of meaning comprehension, "context" would ideally take into consideration factors like experimental participants' subjective moods at the time of testing, aspects of world knowledge, their state of alertness, their language exposure and history, or time of day, week, semester or year when testing was being conducted, to name just a few. While experimenters do their best to hold as many variables as possible constant, individual differences (as well as the fact that multiple ERP subjects cannot usually be tested simultaneously) are, on many fronts, out of their control. Individual participant variation on these and other factors, however, offers yet another potential source for exploring and exploiting data, though beyond the typical screening of language ERP study participants, relatively few experiments collect or analyze results according to such individual differences.

6.12. A look around and ahead

In conclusion, this thesis has offered evidence for implicit, probabilistic, rate-generalizable, left-hemisphere biased anticipatory language processing, which we have argued can be cost-incurring when continuations are highly anticipated but not received. These findings stand in contrast to a more classical view of language comprehension, and brain processing in general, as being essentially bottom-up, waiting for sensory input which is processed and eventually recruited for action. Our findings are more compatible with a new wave of research proposing that a unifying principle for brain operation is one of a more “proactive”, “prospective” or “pre-experiencing” brain (Bar, 2007; Schacter, Addis & Buckner, 2007; Gilbert & Wilson, 2007). Under active brain accounts like these, the brain is assumed to constantly be predicting upcoming input and monitoring the consistency of the anticipated and actual outcomes. This default mode of operation is proposed to occur across domains, at sensory, motor and cognitive levels. With respect to a more cognitive domain, Schacter & Addis (2007) have proposed that a crucial function of memory is to make information available for simulating future events. Under this model, it is unclear what the exact role of a semantic component is in constructing future scenarios, however it seems that prediction in language processing fits nicely with models of predicting upcoming language input based on our stored mental representations in combination with contextual factors. And co-opting another idea from vision research (Ens & Llaras, 2008), it seems possible that our constraint violation finding may be a product of processing information that is inconsistent with some prevailing, preactivated schema or expectation; in turn, the information triggering such discrepancies may ultimately be processed relatively slowly because the parser has to start over restructuring a new contextual representation. These few examples highlight our belief that as we "look ahead" to continued exploration of prediction issues in the language domain, it will also be beneficial to "look around" and let our research be informed, shaped, and spurred by examinations within a larger framework of general brain processing, incorporating proposals of prediction from theories of human motor control (so-called "forward models"), from a variety of aspects of vision research, from judgment and decision making, and from episodic and semantic memory studies; indeed, scientists within these various domains are already doing just this! Without denying the uniqueness and seeming specialization of the human brain for comprehending and producing language, it seems to this researcher that the door has been cracked wider for investigations of how predictive linguistic processing might better be understood in terms of how the brain more generally predicts. With clear evidence *that* the language comprehension system anticipates, we must continue to focus resources on addressing more fine-grained questions of *how* this preactivation occurs.

6.13. References

Altmann, G.T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Anderson, M.C., Ochsner, K.N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S.W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, 303(5655), 232-235.

Aron, A.R., Robbins, T.W. & Poldrack, R.A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8(4), 170-177.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280-289.

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In M. Beeman, & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. (255-284). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Boland, J.E. (2005). Visual arguments. *Cognition*, 95(3), 237-274.

Carlsson, K., Petrovic, P., Skare, S., Petersson, K.M. & Ingvar, M. (2000). Tickling expectations: Neural processing in anticipation of a sensory stimulus. *Journal of Cognitive Neuroscience*, 12(4), 691-703.

Crocker, M.W. & Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, 29(6), 647-669.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science: A Multidisciplinary Journal*, 14(2), 179-211.

Elman, J.L. (1995). Language as a dynamical system. In R. F. Port, & T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition*. (195-225). Cambridge, MA, US: The MIT Press.

Federmeier, K.D. & Kutas, M. (1999). A Rose by Any Other Name: Long-term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41, 469-495.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505.

Gazzaniga, M.S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain*, 123(7), 1293-1326.

Gilbert, D.T. & Wilson, T.D. (2007). Prospection: Experiencing the future. *Science*, 317(5843), 1351-1354.

Gough, P.B., Alford, J.A. & Holley-Wilcox, P. (1981). Words and contexts. In O. J. L. Tzeng & H. Singer (Eds.) (Ed.), Perception of print: Reading research in experimental psychology (85-102) Hillsdale, NJ: Erlbaum.

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, 32(2), 101-123.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science: A Multidisciplinary Journal*, 20(2), 137-194.

Kamide, Y., Altmann, G.T.M. & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133-156.

Knoeferle, P., Crocker, M.W., Scheepers, C. & Pickering, M.J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, 95(1), 95-127.

Kolk, H.H.J., Chwilla, D.J., van Herten, M. & Oor, P.J.W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1-36.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161-163.

Kutas, M., Lindamood, T.E. & Hillyard, S.A. (1984). Word expectancy and event-related potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126-1177.

Mitchell, D.C. (1982). *The process of reading: a cognitive analysis of fluent reading and learning to read*. Published Chichester; New York: Wiley, 1982.

Morris, R.K. (2006) Lexical processing and sentence context effects. In M.J. Traxler & M.A. Gernsbacher (Eds) *Handbook of Psycholinguistics* (2nd Ed.). Academic Press: London.

Novick, J.M., Trueswell, J.C. & Thompson-Schill, S.L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience*, 5(3), 263-281.

Rayner, K., Ashby, J., Pollatsek, A. & Reichle, E.D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 720-732.

Rubia, K., Smith, A. B., Brammer, M.J. & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *NeuroImage*, 20(1), 351-358.

Schacter, D.L. & Addis, D.R. (2007). On the constructive episodic simulation of past and future events. *Behavioral and Brain Sciences*, 30(3), 331-332.

Schacter, D.L., Addis, D.R. & Buckner, R.L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, 8(9), 657-661.

Seidenberg, M.S. & MacDonald, M.C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science: A Multidisciplinary Journal*, 23(4), 569-588.

Simmons, W.K., Martin, A. & Barsalou, L.W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex*, 15(10), 1602-1608.

Slotnick, S.D., Thompson, W.L. & Kosslyn, S.M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex*, 15(10), 1570-1583.

Smith, N.J. & Levy, R. Probabilistic Prediction and the Continuity of Language Comprehension (October 18, 2008). Conceptual Structure, Discourse and Language, 9th Conference on Conceptual Structure, Discourse, and Language (CSDL9).

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632-1634.

Thompson-Schill, S.L. (2005). Dissecting the language organ: A new look at the role of Broca's area in language processing. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones*. (173-189). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V. & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.

van Herten, M., Chwilla, D.J. & Kolk, H.H.J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, 18(7), 1181-1197.

Vissers, C.T.W.M., Kolk, H.H.J., van de Meerendonk, N. & Chwilla, D. J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, 46(4), 967-982.

Wicha, N.Y.Y., Moreno, E. & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*. Vol 16(7), 1272-1288.

Wicha, N. Y., Moreno, E. M. & Kutas, M. (2003a). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483-508.

Wicha, N.Y.Y., Bates, E., Moreno, E.M. & Kutas, M. (2003b). Potatoes not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346, 165-168.

Wlotko, E. W. & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001-3014.

Ye, Z. & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research*, 1203(8 Apr), 103-115.