

ANTICIPATORY LANGUAGE PROCESSING: DIRECT PRE-TARGET EVIDENCE FROM EVENT-RELATED BRAIN POTENTIALS

by

CHRISTOPHER HAMILL

B.A., The George Washington University, 2009

A thesis submitted to
the Faculty of the Graduate School of
the University of Colorado in partial fulfillment of
the requirement for the degree of
Masters of Arts
Department of Linguistics
2012

This thesis entitled:
“Anticipatory language processing:
Direct pre-target evidence from event-related brain potentials”
has been approved by the Department of Linguistics

Dr. Bhuvana Narasimhan

Dr. Laura Michaelis-Cummings

Date: _____

The final version of this thesis has been examined by the signatories,
and we find that both the content and the form meet acceptable presentation
standards of scholarly work in the above mentioned discipline.

IRB protocol #: _____ 0607.15 _____

Hamill, Christopher Owens (M.A., Linguistics, 2012)

Anticipatory language processing: Direct pre-target evidence from event-related brain potentials

Thesis directed by Assistant Professor Bhuvana Narasimhan

This study attempts a novel identification of event-related brain potential (ERP) components of anticipatory processing of semantic information during online language comprehension directly in the pre-target EEG, as well as the directionality of these components (i.e., whether the component's amplitude either increases or decreases with increased semantic constraint). I have done this by manipulating pre-target sentential context constraint across three cloze probability conditions ('high', 'mid', 'low' – Taylor, 1953), and tested for any relation between the pre-target EEG and post-target N400 effects that resulted from the cloze manipulations. Following repeated measures ANOVA and paired t-tests, statistically significant differences were in fact found between the conditions in the pre-target region, and these differences were ordered inversely of cloze probability (i.e., as cloze probability increased, effect size decreased). These results support the increasingly accepted view that the language processor is actively engaged in the predictive processing of future elements of an utterance. Moreover, these data also suggest that semantic information may be processed more rapidly than was originally thought to be the case. Finally, implications for interpreting the N400 effect are discussed.

Contents

I. Introduction

1. The rapidity of language processing	1
2. Anticipatory processing: Definition	4
3. Anticipatory processing: Experimental support	6
4. Methodological assumptions of previous studies	9
5. My study	10
a. Introduction	10
b. Theoretical assumptions	11
c. Hypotheses	13

II. Experimental method

1. Participants	14
2. Materials	14
a. Context norming	14
b. Cloze condition assignment	15
c. Sentence ending assignment and norming	16
3. Procedure	18
a. Experimental procedure	18
b. EEG recording	19
c. Analysis	20

III. Results

1. Behavioral data	22
2. ERP data	23

a. Post-target N400 data	23
b. Target–2	25
c. Target–1	27
IV. Discussion	
1. Direct pre-target evidence of anticipatory processing	29
a. Anticipatory activation and rapid semantic access	29
b. Temporal localization of the pre-target effect	31
c. Post-Target–2 activity	33
2. Directionality of the effect	34
a. Ordering of cloze conditions in the pre-target region	34
b. Relationship between N400 size and the pre-target effect	36
3. Reinterpretation of the N400 effect	36
V. Drawbacks and limitations	
1. Low number of subjects	38
2. Insufficient control of Target–2 and Target–1	39
3. Mean context length between cloze conditions	40
4. Underspecification of findings	41
VI. Directions for future research	41
VII. Conclusions	43
References	45
Appendix	53

Tables

1. Cloze probability and length information for 180 experimental contexts	.	15
2. Frequency and length information for 180 experimental completions	.	18
3. Paired t-test results for the N400 time window	.	23
4. Paired t-test results for the Target-2 time window	.	26
5. Paired t-test results for the Target-1 time window	.	27

Figures

1. ROI electrode sites	20
2. High anomaly vs. high control for the post-target N400 time window	24
3. Mid anomaly vs. mid control for the post-target N400 time window	24
4. Low anomaly vs. low control for the post-target N400 time window	25
5. Mean N400 voltages of semantic anomaly vs. control for all conditions	25
6. Mean voltage of cloze conditions following Target-2 (140-340 ms)	26
7. Pre-target EEG	26
8. Scalp distribution of t-scores at post-Target-1 component peak (200-240 ms)	27
9. Scalp distribution of anticipatory effect strength following Target-1 (190-290 ms)	28
10. Mean voltage of cloze conditions following Target-1 (190-290 ms)	28
11. Graphical representation of spreading deactivation of a semantic network	35

I. Introduction

1. The rapidity of language processing

Of all of the operations regulated by the brain, language processing is among one of the most computationally complex. It requires the simultaneous processing and integration of multiple different kinds of information (i.e., phonological, lexical-semantic, morphosyntactic, pragmatic) and at many different levels of hierarchical and discursive structure, much of which is fluid and constantly being updated and refined by unfolding contextual factors. Considering this fact about language, studies have suggested remarkably that the various features of utterances, irrespective of complexity, are generally processed beyond the level of raw sensory input within just a few hundred milliseconds.

There is much evidence to support this fact about the rapidity of language comprehension. For instance, many studies have investigated early event-related brain potentials (ERPs)¹ of language processing as they apply to morphosyntactic or otherwise non-meaningful linguistic processing. Dikker et al. (2009) manipulated words for which their syntactic category was either overtly marked by a closed-class function morpheme or not marked at all. They found that when expectations about the target word's syntactic class were violated, this manipulation affected processing as early as 125 ms after stimulus onset in the form of an increased M100 effect over occipital regions, suggesting that syntactic cues are being processed at least this early (for a similar study with similar

¹ For those unfamiliar with event-related potentials, ERPs (or 'ERP components') refer to any measured neural activity resulting from cognitive processing. ERPs come in the form of voltage spikes (either positive or negative) of electrophysiological activity that can be reliably correlated with some manipulation of a particular psychological task. Some of the more well-known ERPs for language processing include the N400 (Kutas & Hillyard, 1980), the P600 (Osterhout & Holcomb, 1992), and the early left anterior negativity (ELAN) (Friederici et al., 1993).

results see Dikker et al., 2010). Similarly, Neville et al. (1992) reported early ERP differences as a function of this same linguistically meaningful word class distinction (i.e., open vs. closed). By 150 ms after stimulus onset, open class words had elicited significantly higher P200 effects than closed class words over the left hemisphere, and significantly larger N170 effects over posterior and occipital electrode sites. Pulvermüller et al. (1995) discussed the same finding as Neville et al. (1992), with a substantial open/closed class ERP distinction by 160 ms post-stimulus onset, and potentials for closed class function words being generally more negative over the left hemisphere than the right. Neville et al. (1991) reported evidence of syntactic processing effects around 125 ms, by which point phrase structure violations (e.g., ‘...Max’s of proof the theorem’, ‘...Ted’s about films America’) resulted in a slow, sustained negativity over anterior regions of the left hemisphere. Lau et al. (2006) found evidence that in highly syntactically constraining situations, structural processing expectations are in place by at least 200 ms, long before the P600 (i.e., 600 ms after stimulus onset) which has traditionally been used to infer syntactic processing following Osterhout & Holcomb (1992) (see Section I.4 below for a brief discussion of the P600 and other language-related ERP components).

Elements of meaningful semantic processing have also been shown to occur on a similarly short timescale. Dehaene (1995) reported bilateral posterior brain responses that differed between meaningful linguistic symbols (i.e., animal names, verbs, numerals, and proper names) and meaningless consonant strings by 192 ms following stimulus onset, with non-significant left lateralization beginning to appear between these conditions as early as 138 ms. This finding strongly suggests a very rapid identification of these items

along linguistic dimensions. Pulvermüller et al. (2001) presented a single subject with classes of nouns that had either unimodal (visual) or bimodal (visual and action) associations, and measured the neuromagnetic fields associated with the visual presentation of each class. The bimodal nouns evoked stronger magnetic fields than unimodal nouns by 100 ms, providing evidence of semantic access for visually presented words which, according to the authors, is earlier than any study which reports rapid syntactic access. Moreover, further differences were observed in the neural signal between nouns, verbs, and function words over the left hemisphere by 150 ms. By manipulating the degree of semantic association between the target and certain words in the pre-target sentential context, Penolazzi et al. (2007) demonstrated that cortical sensitivity to semantic context can be observed using electroencephalography (EEG) by 200 ms. Using reading times and an acceptability judgment task, Duffy (1986) found that subjects were slower to read and judge target sentences as unrelated when they followed high expectation-generating texts, and vice-versa. As with the previous studies just introduced, these findings suggest that subjects had been progressively building up selective and context-specific expectations about what may follow based on what had been already encountered.

Thus, a rich literature underlies the now widely held assumption that certain aspects of language processing may be much faster than originally conceived. The question can be posed then, how is this remarkable speed possible given the complexity of the computation and integration involved? Historically models of language processing have been strictly serial in nature, where the various kinds of information involved in understanding an utterance are processed sequentially rather than in parallel (see the

following for some better known serial models of language processing: Friederici, 2002; Geschwind, 1967; Indefrey & Levelt, 2004; Price, 2000; for an overview see Ben Shalom & Poeppel, 2008). Moreover, many of these models posit that syntactic information is categorically processed first, and that it is from the output of this syntactic processing that semantic and pragmatic factors are then computed.

A growing body of research has begun to suggest however that serial models of language processing as traditionally conceived may be insufficient to handle the rapidity of language, and at worst that these models may be simply wrong. As a response, some psycholinguists and cognitive neuroscientists have offered a different possibility that is better equipped to explain at least the speed of language comprehension: anticipatory processing (e.g., Bar, 2007; Bendixen et al., 2012; Federmeier, 2007).

2. Anticipatory processing: Definition

‘Anticipatory processing’ (also variously called ‘prediction’ or ‘predictive processing’ in the literature – Bendixen et al., 2012; Federmeier, 2007; Frisson et al., 2005; Kamide et al., 2003a; Kamide, et al., 2003b) here refers to the prediction of upcoming elements based on those that have already been explicitly uttered. This anticipatory processing can come in a number of forms. Depending upon the nature and degree of constraint introduced by the prior linguistic context, a number of features of the succeeding utterance can be predicted by the language processor. Anticipatory processing can employ linguistic or encyclopedic information to predict, for instance, the

grammatical class of an upcoming item, the thematic role of a future argument in a predication, or the temporal relation of an action², to name a few.

What this predictive process essentially amounts to for language processing is the pre-activation of cognitive and linguistic structures in the brain which facilitates their speedy integration into the larger discourse if and when they are actually encountered. Because these elements are preemptively activated prior to their actual encounter, in essence the resting activation level for these structures is raised, and it will thus take net less cognitive effort to integrate them into the larger discourse when the time comes.

When phrased this way, the concept of anticipatory activation is hardly new. The existence of this form of preemptive activation has long been observed to occur in motor cortex prior to physically initiating a motor action; the logic just described above applies here as well, in that increased resting activation levels require less additional activation in the future when the anticipated motor sequence actually needs to be initiated (Aruin & Latash, 1995; Tanji & Evarts, 1976). In fact, not only is this anticipatory activation likely to occur prior to motor movements, but in fact a heightened ability to anticipatorily activate task-relevant motor structures may actually be correlated with motor movement expertise (Aglioti et al., 2008). The same kind of motor pre-activation has been robustly observed to occur during fluent speech across languages as well, and has been termed the ‘anticipatory coarticulation’ of articulatory gestures (Fowler & Saltzman, 1993; Guenther, 1995). As a speaker produces a certain sound in a word, the active articulators (e.g., lips, tongue) are constantly moving preemptively into position to rapidly pronounce the following speech sound when the preceding segment is complete (Gandour et al., 1994;

² See (Langacker, 1987) for an interesting discussion of how temporal relations in language can be explained with exclusive reference to cognitive/semantic representations.

Gibbon et al., 1993; Lubker & Gay, 1982; Strange & Bohn, 1998); an analogous phenomenon has even been identified for sign language (Jerde et al., 2003). Both anticipatory coarticulation and non-speech-related muscular pre-activation result in a situation where subsequent task-relevant actions can be initiated more quickly than would have been the case if the relevant structures had remained dormant until needed. An analogous claim is being made here regarding the cause and effect of anticipatory activation in the brain during rapid language processing.

There is some debate as to the source of anticipatory activation as it applies to language. In particular, opinions differ as to whether the kind of pre-activation discussed here stems only from the semantic/syntactic priming of a target by specific words in the preceding sentential context, or if pre-activation can also result from the understanding and integration of higher order linguistic structures such as those that build throughout an entire sentence or discourse. While there is no doubt that local lexical priming occurs during natural language processing, recent electrophysiological evidence has surfaced to suggest contrary to the accepted wisdom that anticipation can be driven by communicative elements at other levels as well, such as global discursive context (Nieuwland & Van Berkum, 2006; Otten & Van Berkum, 2008; Van Berkum et al., 1999). Anticipatory processing can thus be thought to occur at many different levels of language use, and its importance for sentence processing is only just beginning to be understood.

3. Anticipatory processing: Experimental support

A large amount of experimental support exists to back up claims about the reality of anticipatory processing during language comprehension. For instance, by exploiting

the phonological regularity of the indefinite articles ‘a’ and ‘an’ in English, DeLong et al. (2005) found that the morphophonological form of a pre-target word can influence a listener’s expectations for which specific item will follow, independent of the target’s semantics or contextual appropriateness. In fact, DeLong et al. reported finding that online N400 effect size decreased as a target word’s offline cloze probability (Taylor, 1953) increased, suggesting that higher predictability is correlated with lower N400 amplitude. Similarly, a variety of other studies using languages other than English have shown that listeners’ (or readers’, depending on the modality) expectations for an upcoming noun’s grammatical gender given a certain context can modulate ERP effects, suggesting that the noun’s gender had been activated before the noun itself was actually encountered (Van Berkum et al., 2005; Wicha et al., 2003a; Wicha et al., 2003b; Wicha et al., 2004). Dambacher et al. (2009) reported evidence of very early-latency interactions (90 ms) between prediction-driven top-down lexical semantic processing (i.e., expectations supplied by the language processor for what semantic content is yet to occur) and input-driven word form analysis (i.e., bottom-up processing of basic level sensory inputs that have already been encountered), further suggesting that anticipatory activation can influence the processing of early exogenous inputs.

Altmann & Kamide (1999) presented evidence from the eye-tracking paradigm that semantic and syntactic valency information introduced by the verb is used in an anticipatory manner to restrict the possible future arguments of that verb. Similarly, Altmann & Kamide (2007) showed with another eye-tracking experiment that contextual cues and world knowledge about objects (which they refer to as an object’s ‘affordances’) are integrated with syntactic information (in this case, verb tense) to help disambiguate

possible continuations of an utterance. Kamide et al. (2003a) and Kamide et al. (2003b) offer additional evidence from a handful of ‘visual-world’ experiments that the integration of morphosyntactic case marking on preverbal nominal arguments, the semantic constraints of the verb, and real-world knowledge can result in the ability to actively predict elements of an utterance downstream of the verb. Using pairs of visual (Federmeier & Kutas, 1999) and auditory stimuli (Federmeier & Kutas, 2002), Federmeier & Kutas demonstrated that post-target N400 amplitude is modulated by the degree of semantic constraint of the pre-target context. They found that when a context primed a specific target exemplar, targets with partial semantic category overlap with the primed exemplar were processed more easily (as indexed by lower – that is, less negative – N400 amplitudes) than those with no overlap at all, with the exemplar itself yielding the lowest N400 effect. This is what one would expect to find if the language processor actively anticipates elements of an upcoming target based on previously encountered material.

Thus the claim that the anticipatory activation of upcoming structures occurs during sentence comprehension appears now to be robustly supported by experimental evidence from multiple methodologies. Moreover, this predictive mechanism does not appear to be limited to processing only a single type of information, but rather occurs simultaneously in the phonological, morphosyntactic, semantic, and discourse-level domains. With this as the intellectual backdrop, in order to lay the groundwork for the study introduced in this paper, it will now be helpful to elucidate some of the methodological assumptions that many of these studies hold in common, and which will be exercised in the interpretation of my results as well.

4. Methodological assumptions of previous studies

Most if not all of the studies that have been conducted up until the present time regarding event-related brain potentials and language processing have relied heavily on certain physiological assumptions regarding the functional significance of various ERP components. For instance, the mismatch negativity (MMN) has been taken to be a signal of the recognition of the physical abnormality of a stimulus compared to other habituated stimuli (Näätänen et al., 1978); from this it can then be inferred that there must be some sort of pattern recognition mechanism via which a deviant stimulus is identified, thus resulting in an MMN. Another component, the N400, is generally seen as an index of semantic anomaly detection (Kutas & Hillyard, 1980), from the presence of which it can then be inferred that semantic expectations had been built up prior to encountering the anomaly. The P600 is yet another component which is most often taken to be a signal of syntactic or otherwise structural processing (Osterhout, 1992) from which, not unlike the MMN or the N400, it can be inferred that structural predictions had been generated and subsequently violated, resulting in the P600 component.

With these commonly held interpretations in mind, these observable ERP components and others are used by cognitive neuroscientists to make inferences about the unobservable cognitive processes that underlie or give rise to ERPs. Scientists therefore tacitly assume that the manifestation of ERPs signifies the existence of certain cognitive processes under investigation. This may seem like a tautology and I do not wish to challenge such assumptions here, however it is an important belief about the nature of ERPs that needs to be made explicit, and which makes possible the methodology of the experiment introduced in this paper.

5. My study

a. Introduction

As just discussed, all previous ERP research has used the presence or absence of various ERP components to infer facts about the nature of unobservable cognitive processes occurring ‘behind the scenes’ in the mind. This is particularly true in the case of studies looking at anticipatory processing, which have traditionally only analyzed post-target neural activity in order to infer aspects of the pre-target predictive mechanism. This is an extremely helpful way to approach the issue and has certainly not been without results. However, because the anticipatory processing crucially occurs prior to the presentation of the target, investigations of post-target activity alone are insufficient to fully uncover the nature of the anticipatory language processing mechanism.

Consequently, my study makes a novel attempt to directly observe the anticipatory processing mechanism while this online prediction is actually occurring, rather than exclusively measuring neuro-electrical signals after the target and making inferences about pre-target activity on the basis of that alone. In this respect, the study introduced here is truly the first of its kind, and represents an exploratory analysis into the pre-target region that the current literature simply does not address. In order to do this, I will measure pre-target EEG activity³ directly in order to make a cursory attempt to identify what this anticipatory activation looks like in real time EEG. My goals are two: (1) To identify a hitherto undiscovered effect in the pre-target EEG that results from linguistic prediction alone; (2) To determine the ‘directionality’ of this effect; that is, the

³ Roughly, ‘EEG activity’ (or just ‘EEG’ for short) refers to the sum electrophysiological output of the brain at a particular time. EEG does not reference any particular functional component, such as an event-related potential.

correlation between effect amplitude and degree of anticipatory processing (e.g., Do higher amplitudes correlate with more prediction or less prediction?).

b. Theoretical assumptions

To accomplish these goals, my study will require a chain of theoretical assumptions that are supported by the extant literature. First, I will assume that the presence of an N400 effect in my data reflects the encountering of a stimulus that was semantically unexpected given a previous context (Assumption 1). This assumption is in keeping with both the methodological assumptions and the findings of many previous studies (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000; Kutas & Federmeier, 2011). My stimuli sentences have therefore been engineered with Assumption 1 in mind such that each sentence comes in two varieties: one ending in a word that is semantically incongruous with the preceding context ('context' here referring to every word of a sentence except the final one, the target), and a semantically well-formed control. The semantically anomalous version is expected to elicit a robust post-target N400 effect relative to the control, and in following the assumptions of previous studies such as those just cited, these effects will be interpreted as reflecting the 'poorness of fit' of the sentence-final target with the predictions anticipatorily generated by the language processor.

Second and subtly different from Assumption 1, I will assume from the presence of post-target N400 effects that semantic predictions for the sentence-final word had in fact been generated in advance on the basis on prior contextual cues (Assumption 2). As just mentioned, the N400 is assumed to reflect the semantic unexpectedness of a stimulus.

It follows thus that in order for a stimulus to be deemed unexpected, there must have been a sufficiently specific expectation to begin with. This, in a word, is the anticipatory processing mechanism that I am interested in examining here.

Third, I will assume that the amplitude of the N400 effect is modulated by the degree of semantic deviation from given predictions (Assumption 3). That is to say, the higher the degree of deviation from the expectations of the language processor, the larger the N400 effect will be. This assumption is well supported by previous research, which has found that N400 effects are the smallest following completely primed entities, higher following partially primed entities, and higher still for completely unprimed (i.e., unrelated) entities (Federmeier & Kutas, 1999; Federmeier & Kutas, 2002; Pratarelli, 1995). This fact about the N400 will help characterize the directionality of the anticipatory processing effect I hope to identify.

Due to the inherent difficulty of quantifying semantic anomaly, rather than design sentence completions of increasing contextual violation, in my experiment I have manipulated context constraint in the form of offline cloze probability (Taylor, 1953) in order to control the degree of semantic prediction occurring for any particular target word. ‘Context constraint’ refers to the range of possible completions of a sentence as established by norming studies (see Section II.2.a-b below for the details of this norming process), where only one or two acceptable completions is considered high constraint (see Schwanenflugel & Shoben (1985) and Schwanenflugel & LaCount (1988) for a more in-depth discussion of context constraint). Therefore, I have engineered three experimental conditions that differ only in their degree of context constraint, and this

three-way distinction should allow me to observe the directionality of any shifts in EEG stemming from the degree of anticipatory processing.

c. Hypotheses

I have two specific hypotheses regarding what my data will reveal. Hypothesis 1, following from Assumptions 1 and 3 above and consistent with the findings of a number of previous studies (Bentin et al., 1985; DeLong et al., 2005; Federmeier, 2007; Kutas & Hillyard, 1984; St. George et al., 1997; Van Berkum et al., 1999), is that N400 amplitude to semantic anomaly should be directly proportional to an item's offline cloze probability (i.e., its context-dependent predictability), such that one should be roughly predictable given the other. More explicitly, the semantically anomalous versions of high cloze stimuli should yield the largest N400 effect, followed by mid and then low cloze stimuli. This is because items with a high cloze probability are highly expected, and analogously to simpler word-pair experiments where the first word primes the second, providing a predicted form should yield a low N400 relative to a less predictable or expected item.

Hypothesis 2, pending the robust identification of a cloze-correlatable pre-target anticipatory effect, is that the size of this effect should be inversely proportional to the amplitude of the post-target N400. By using my three-way-condition experimental design and employing the three theoretical assumptions about the significance and dynamics of the N400 effect discussed above, I should be able to directly identify some aspect of the directionality of the anticipatory processing as it manifests itself in EEG.

II. Experimental method

1. Participants

17 subjects were recruited for this study between the ages of 18 and 34 (mean age = 22.4; 7 females). Subjects were native speakers of American English. All had normal or corrected-to-normal vision, no known history of neurological impairment, and were right-handed. Participation was voluntary, and both verbal and written consent was obtained prior to beginning the experiment.

2. Materials

a. Context norming

417 sentences were either created or borrowed from previous studies (Block & Baldwin, 2010; Bloom & Fischler, 1980; Hamberger et al., 1996; Kim & Lai, In Press). The final word was removed from each sentence, and either 30 or 40 anonymous volunteers (depending arbitrarily on the sentence) were asked on Amazon Mechanical Turk to provide the most appropriate completion for each. The workers were all native speakers of English located in the United States, and had to pass a number of simple tests in order to determine that they were not bots (i.e., software programs designed to imitate the behavior of real people online). Acceptable completions could be of any grammatical class, provided that they were appropriate to the sentence context. From this database, sentences with plurality completions⁴ of 75% or higher, 37.5-62.5%, or 25% or lower were selected to comprise the three cloze conditions, discussed in more detail below.

⁴ Given a set of appropriate completions provided by multiple people for one sentence, that sentence's 'plurality completion' refers to the most common completion in the set, regardless of whether this word accounts for more or less than 50% of the total number in that set.

b. Cloze condition assignment

After norming on Amazon Mechanical Turk, 327 sentences remained that would be satisfactory for this study given the cloze condition requirements. Of these 327 sentences, 180 were selected for use in this experiment (see Appendix for the complete list). These 180 sentences were divided into three groups of 60 ('high', 'mid', and 'low') defined by the cloze probability of their members. The 'high cloze' group was composed of sentences with cloze probabilities of 75% or more; the 'mid cloze' group was composed of sentences with cloze probabilities between 37.5% and 62.5%; the 'low cloze' group was composed of sentences with cloze probabilities of 25% or less. The average length (defined as number of words minus the target completion, addressed in the following section) was kept as constant for each cloze condition as the prefabricated⁵ stimuli would allow (see Section V for a discussion of some potential problems with this prefabrication). This information is summarized in Table 1 below.

Cloze condition	Cloze probability: mean, range	Length (no. of words): mean, range	Example
High (≥75%)	99.2%, 97.6-100%	10.5, 6-19	The cowboy put a saddle on the ...
Mid (37.5-62.5%)	53%, 46.7-62.5%	8.3, 4-17	The squirrel scurried away from the ...
Low (≤25%)	19%, 6.7-25%	7.3, 5-13	There was nothing wrong with the ...

Table 1. Cloze probability and length information for the 180 experimental sentences before the addition of sentence-final completions.

⁵ The experimental stimulus contexts are described here as 'prefabricated' because the majority of them were inherited from previous studies (Bloom & Fischler, 1980; Hamberger et al., 1996; Block & Baldwin, 2010; Kim & Lai, In Press). Because these previous studies did not all have the same research objectives in mind, their cloze-normed sentences were not all controlled according to the same criteria (e.g., number of words). This admittedly made it quite difficult to perfectly control for this variable for the study reported here. This issue is addressed in more detail in Section V below.

In addition to the 180 experimental stimuli, 60 well-formed filler sentences were also included to obscure my design from the subjects. Each of these filler sentences was followed by a True/False comprehension question, the purpose of which was twofold: (1) To provide an additional criterion apart from raw EEG quality to determine whether or not subjects provided high quality data and should be included in the final analysis; (2) To force subjects to actively attend to the stimuli throughout the entire recording session. (2) was ensured by dispersing the filler sentences randomly among the experimental items so that subjects could not predict when a comprehension question would be presented next. These questions were worded as declarative statements to which subjects replied either ‘True’ (i.e., ‘Given the previous sentence, this statement is most likely true.’) or ‘False’ (i.e., ‘Given the previous sentence, this statement is most likely false.’). Because there was often not a single unequivocally correct answer, in order to respond appropriately to these comprehension questions, subjects were required to make non-trivially difficult inferences from the content of the preceding filler sentence in order to provide accurate responses (e.g., Sent.: ‘Bill jumped in the lake and made a big splash.’ CompQ.: ‘Bill didn’t know how to swim.’ Ans.: ‘False’ – because we can assume that if Bill did not know how to swim, he would not have jumped into the lake in the first place).

c. Sentence ending assignment and norming

Once this cloze condition assignment was complete, each sentence was given a semantically anomalous ending in addition to the well-formed control ending that had been provided by the workers on Amazon Mechanical Turk. The semantically anomalous endings were chosen by recycling well-formed control endings of sentences that had not

been selected for this study and arbitrarily attaching them to one of the 180 experimental items. I then double-checked each sentence to ensure that none of the pseudo-randomly assigned completions actually resulted in a semantically well-formed control sentence; in these cases, another word was arbitrarily selected. Next, heeding the findings of previous studies that word frequency and length can affect ERP component latency and amplitude (Osterhout et al., 1997; Assadollahi & Pulvermüller, 2001; Assadollahi & Pulvermüller, 2003), the control and semantically anomalous ending pairs were normed for both frequency and length. Frequency norming was accomplished by calculating the frequency of each completion's occurrence in the 425 million-word Corpus of Contemporary American English (COCA). Because several endings were extremely high frequency (e.g., 'her'), natural log transformed frequency measurements were used, which greatly reduce the effect of these outlier words during frequency averaging. Frequency information was calculated first for the well-formed control completions, as these completions were provided by the cloze norming process on Amazon Mechanical Turk described above and thus could not be modified. The same process was then undertaken for the semantically anomalous completions, and when the resulting mean natural log transformed frequencies was not matched across anomaly and control endings within a condition, I selected new anomalous endings where necessary to match with the normed well-formed control endings. The final frequency information is summarized in Table 2 below. A similar norming process was undertaken for mean length (defined as number of characters) of the well-formed control and semantically anomalous endings between the three cloze groups. Namely, the length of each unalterable control ending was calculated and then averaged for each condition, followed by the semantically anomalous endings; where additional

tweaking was necessary to match anomaly and control within and across conditions, new anomalous endings were chosen to shift the average in the desired direction.

Cloze condition	Control vs. anomaly	Natural log frequency (COCA): mean, range	Length (no. of characters): mean, range
High	Control	10.2781, 7.9848-12.0893	4.550, 4-5
	Anomaly	10.22251, 7.8782-12.5459	4.683, 3-7
Mid	Control	9.9337, 5.7430-12.3000	4.917, 3-8
	Anomaly	9.9692, 7.8555-13.0819	5.083, 3-10
Low	Control	9.8204, 4.3820-13.8271	5.800, 3-13
	Anomaly	9.9230, 7.1678-12.6420	5.717, 3-9

Table 2. Natural log transformed frequency and length information for the 180 experimental sentence completions.

3. Procedure

a. Experimental procedure

Subjects sat in a comfortable chair in a silent, soundproof booth with the lights off to prevent any sensory distraction from the experimental task. They were positioned approximately 100 cm from a screen that displayed experimental stimuli one word at a time in rapid visual serial presentation (RSVP), each preceded by a fixation cross which lasted for 750 ms. Each word was presented for 200 ms, with a stimulus onset asynchrony (SOA) of 450 ms (stimulus presentation: 200 ms; inter-stimulus interval (ISI): 250 ms). Instructions were standardized so that each subject would receive the same information regarding how to take the experiment. Each subject was then presented with the 180 experimental sentences discussed above plus the 60 filler sentences (total = 240 sentences), broken up into four blocks of 60 sentences each. All subjects saw the same 180 experimental sentence contexts, but not all subjects saw the same endings. The endings of 30 of the 60 sentences within each of the three cloze groups were well-formed (control), while the remaining 30 within each group were semantically anomalous. Thus,

each subject saw a total of 90 well-formed and 90 semantically anomalous sentences, balanced equally across the cloze conditions. These lists of 240 experimental stimuli and filler sentences were initially ordered completely randomly, then after a visual inspection of the resulting randomization, any chunks of three or more consecutive stimulus items of the same condition were broken up by an arbitrarily selected item from a different condition. This effort was taken to prevent subjects from correctly identifying one of the experimental manipulations. Following this pseudo-random ordering of the lists of stimuli, each list was then counterbalanced by presenting half of the subjects with the stimuli in a forward order and half of them in a backwards order to avoid order effects. For their behavioral responses to the comprehension questions following filler sentences, subjects were given a button box that they held on their laps and were instructed to operate with both hands to minimize accidental incorrect button presses. True/False behavioral responses were balanced for both right and left hand bias. Because subjects were in control (via the button box) of when to proceed onto the next stimulus sentence, the experiment was effectively self-paced, which gave subjects the ability to rest their eyes between sentences and only continue onto the next item once they were ready.

b. EEG recording

Continuous EEG was recorded from 64 sintered Ag/Ag-Cl electrodes embedded in a plastic cap (NeuroScan Quik-Caps) arranged according to the extended 10-20 system (see Figure 1 on the following page). Blinks and vertical eye movements were recorded by two electrodes attached above and below the left eye, with horizontal movements recorded by electrodes placed at the outer canthi of each eye. EEG was also recorded

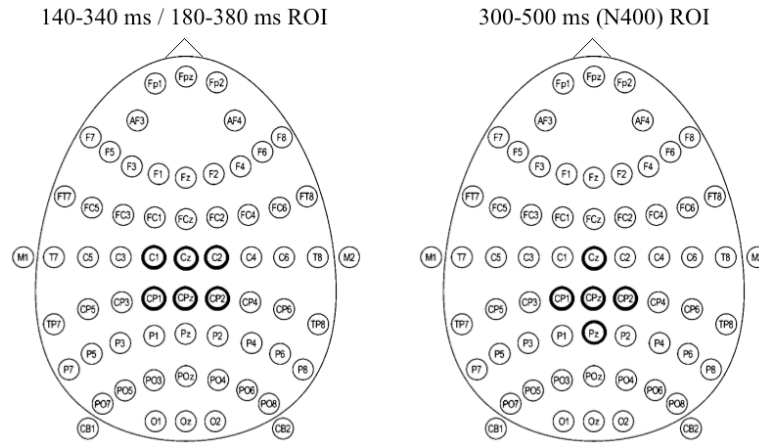


Figure 1. Electrode sites comprising the two ROIs (see Footnote 7 for a definition of ROIs) used for the calculation of ERPs.

referenced offline to linked mastoid channels.

EEG was amplified and digitized at 1000 Hz (NeuroScan Systems). Following the recording, all offline EEG data was decimated to 200 Hz and subjected to a bandpass filter of 0.1-50 Hz. Ocular artifacts were corrected using a subject-specific regression-based algorithm (Semlitsch et al., 1986). Any remaining voltages in excess of $\pm 100 \mu\text{V}$ were rejected. ERPs were averaged across all subjects in epochs of activity spanning -200 ms to 1450 ms relative to the antepenultimate word of each sentence⁶, and -200 ms to 1000 ms relative to the sentence-final target.

c. Analysis

ERPs were time-locked to the presentation of individual words within the stimulus sentences. ERPs were quantified as mean voltages within three time windows,

⁶ Because this study seeks to directly identify an effect of anticipatory processing located in the pre-target region, and because prior literature does not address the issue of where within this region such an effect may occur, the relatively long time window (1650 ms total) was selected for epoching. One hypothesized shape for this yet unidentified pre-target effect is drift occurring over several words, and thus this time window (which encompasses the presentation and immediately subsequent neural activity of the antepenultimate word, the penultimate word and the sentence-final target at a single glance) was deemed most appropriate to view such a multiword effect.

taken from (1) 140-340 ms after the presentation of the antepenultimate word of the sentence (hereafter referred to as Target–2, or ‘target minus two’), (2) 190-290 after the presentation of the penultimate word of the sentence (hereafter referred to as Target–1, or ‘target minus one’), and (3) 300-500 ms after the presentation of the sentence-final target (hypothesized N400 effect). Lacking any previous studies with similar goals to use as a guide for what to expect, the first two time windows (140-340 ms after Target–2 and 190-290 ms after Target–1) were selected on the basis of a visual inspection which consisted of viewing grand averaged data at each electrode individually and attempting to identify where in time the largest, most robust, and most widely-distributed cloze probability-dependent pre-target effect(s) appeared to be located. Upon this identification, the durations of the Target–2 and Target–1 time windows were selected to be the specific slice of time where the effect seemed clearest between the cloze conditions. Although the latency of the post-Target–1 time window is slightly later (+50 ms) than the post-Target–2 window, I assume that the same component is being measured in both cases, as previous studies have shown that the latency of a single component can vary greatly both between and within subjects (Luck, 2005). Thus, the 50 ms discrepancy in the latency of these two time windows is not a source of great theoretical concern. In contrast, the hypothesized N400 time window (300-500 ms after the target) was chosen on the basis of prior N400 research (Deacon, et al., 1995; Kim & Lai, In Press; Kutas & Hillyard, 1980).

Mean voltages within the three time windows were measured from centro-parietal channels collapsed into two separate regions of interest (ROIs)⁷. The ROI used for the

⁷ In neuroimaging studies, an ROI is an area of the brain where “the effects of interest [are] present” (Luck, 2005). Because of the ever-present need in ERP research to average across large amounts of EEG data, ROIs are calculated by averaging the mean activity of several electrodes

measurements of the two pre-target time windows was comprised of the electrode sites C1, CZ, C2, CP1, CPZ, and CP2, while the ROI used for measuring the post-target N400 time window was comprised of electrode sites CZ, CP1, CPZ, CP2, and PZ, as is typical in the literature for N400 analysis; Figure 1 on page 20 shows these specific electrodes.

Subjects who provided behavioral data (i.e., responses to the True/False comprehension questions that followed all filler sentences) with a mean accuracy of below 80% were removed from this study. The rationale for this relatively low level of criterion (many studies set criterion at or above 90%) was that the comprehension questions were purposefully designed to be somewhat difficult, and it was therefore expected that throughout the entire 1.5-2 hour recording session, subjects would inevitably respond incorrectly to several items, even if by accident.

III. Results

1. Behavioral data

The average accuracy of the responses to comprehension questions following filler sentences was 92.39% with a range of 83.33-98.33%; thus the data of all 17 subjects was included in the analyses of Section II.3.c. All but three subjects performed with mean accuracies at or above 90%. No subject performed at 100%, but again this was to be expected due to the difficulty of some of the comprehension questions involved.

within a particular region to arrive at a single central tendency that characterizes the activity of that entire region within a certain span of time.

2. ERP data

a. Post-target N400 data

Analyses were repeated measures analyses of variance (ANOVA) with the factors condition (high cloze, mid cloze, and low cloze) and the difference between semantic anomaly and well-formed control (baseline). Following a 2 (anomalous vs. control) x 3 (high cloze, mid cloze, low cloze) ANOVA of ERP data in the post-target N400 region (i.e., 300-500 ms following the presentation of the sentence-final target), a significant main effect of cloze probability was observed [$F(2, 32) = 8.6, p < 0.05$]. There was also a main effect of semantic anomaly [$F(1, 16) = 34.636, p < 0.05$], revealing the N400 amplitudes to be significantly different between anomaly and control within each

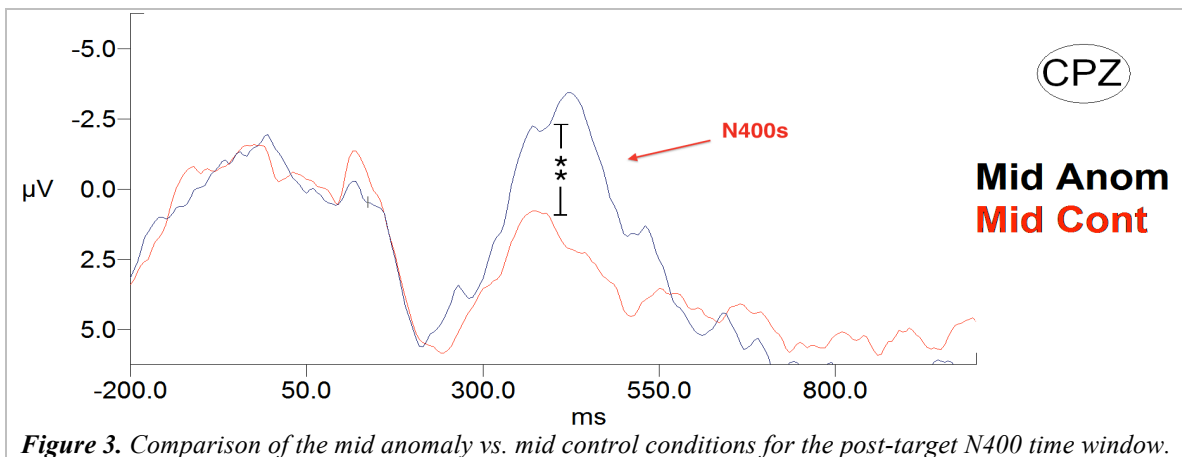
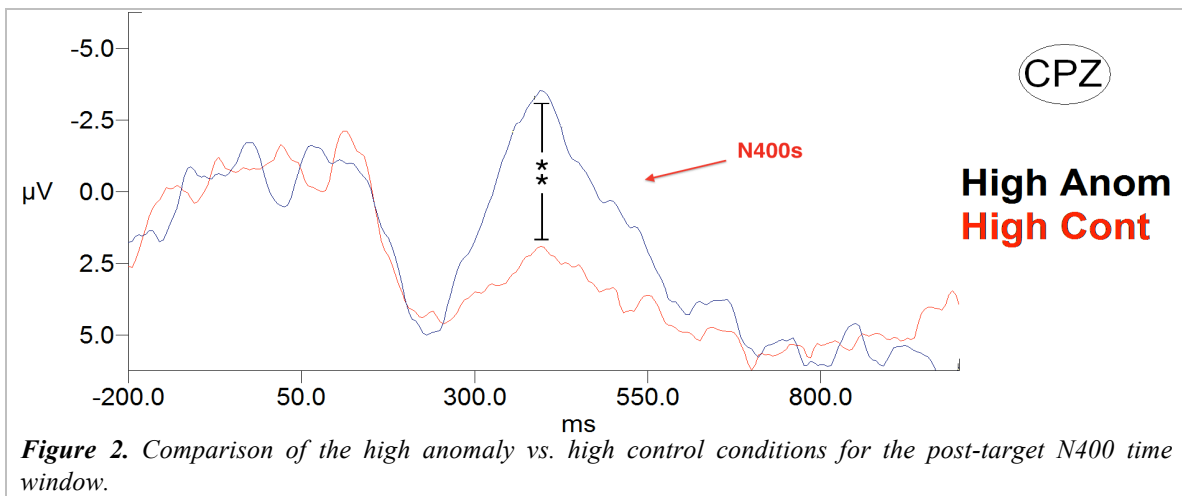
Conditions compared	<i>t</i>	<i>df</i>	<i>p</i>
High-Mid	-1.1404	16	0.2709
High-Low	-1.2998	16	0.2121
Mid-Low	-0.2835	16	0.7804

Table 3. Paired *t*-test results for the N400 time window.

condition, as hypothesized (see Figures 2-5). However, no significant interaction was observed between the cloze conditions and the presence or absence of semantic anomaly [$F(2, 32) = 1.0467, p = 0.3628$]. In response to this lack of significant cloze-by-anomaly interaction, paired *t*-tests were conducted to see if the voltage differed significantly between anomaly and control across any two of the cloze conditions. Surprisingly, no significant differences were found between any of the conditions. The results are summarized in Table 3 above.

Two more important observations can be made regarding this time window. First, the differences in N400 amplitude between anomaly and control elicited across the cloze conditions were correlated with cloze probability. The high cloze condition produced the largest N400, followed by the mid and then the low cloze conditions (see Figures 2-5

below and on the following page). Consequently it can be said in accordance with Hypothesis 1 that N400 effect size is positively correlated with cloze probability. Second, a visual inspection of the data reveals that the peak amplitude of the N400 effect of semantic anomaly appears to remain fairly stable across all cloze conditions, while the amplitude of the well-formed control (baseline) appears to decrease with constraint (see Figures 2-4 for a comparison of the electrode CPZ across the conditions – CPZ will be used as the representative electrode throughout due to its central position in both ROIs, however the patterns observed there are similar at all electrodes that constitute the ROIs). These differences are not significant, but the apparent visual trend is curious and merits further discussion (see Section IV.3).



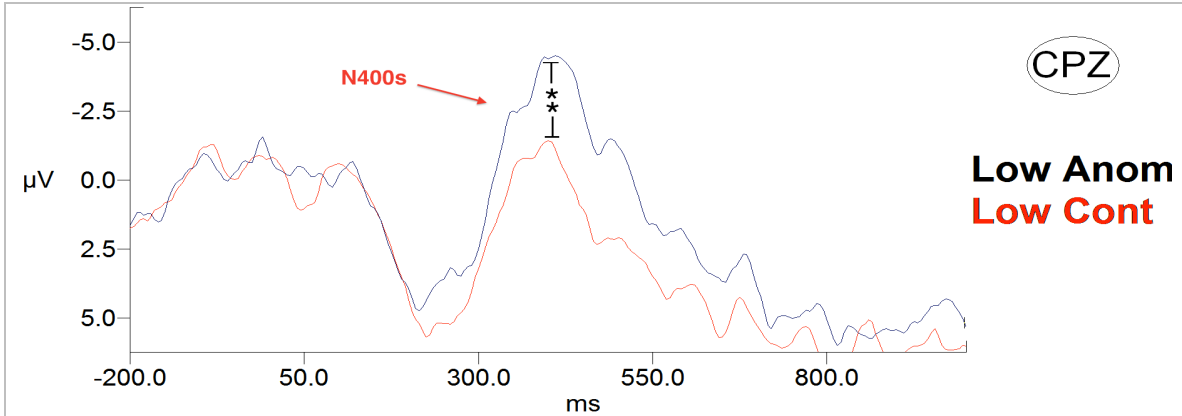


Figure 4. Comparison of the low anomaly vs. low control conditions for the post-target N400 time window.

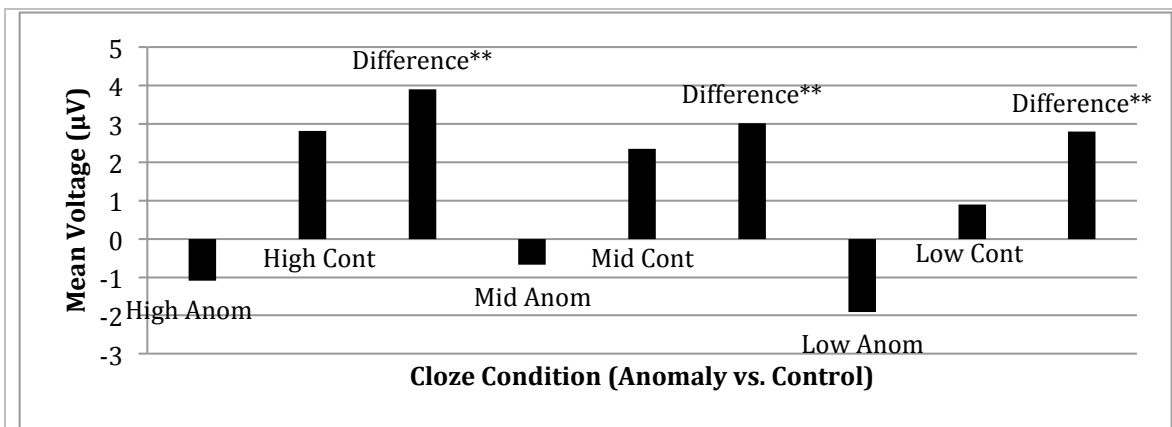


Figure 5. Mean voltages of semantic anomaly vs. control within the cloze conditions, and the differences between them.

b. Target-2

One large positive-going wave was observed following both Target-2 and Target-1 at the time windows specified in Section II.2.c. First, a repeated measures ANOVA of this component (collapsed across the anomalous and control conditions in order to increase power, as the anomalous/well-formed distinction only applied after the target had been presented) following the presentation of Target-2 (140-340 ms) revealed no significant main effect of cloze probability [$F(2, 32) = 1.272, p = 0.2941$]. As with the post-target N400 data just described, following this lack of statistical significance, paired t-tests were conducted to investigate the statistical relationship between the mean vol-

tage of each condition individually (again, collapsed across anomaly and control

Conditions compared	<i>t</i>	<i>df</i>	<i>p</i>
High-Mid	-0.833	16	0.4171
High-Low	0.6263	16	0.54
Mid-Low [†]	-1.9828	16	0.06483

Table 4. Paired *t*-test results for the Target-2 time window⁸.

conditions). Again, none of these means were found to be significantly different, though the difference between the mid and low conditions is quite close for what it is worth. The results are summarized in Table 4 above. See Figures 6 and 7 below for a comparative graph of the means of each condition and a snapshot of the pre-target EEG, respectively.

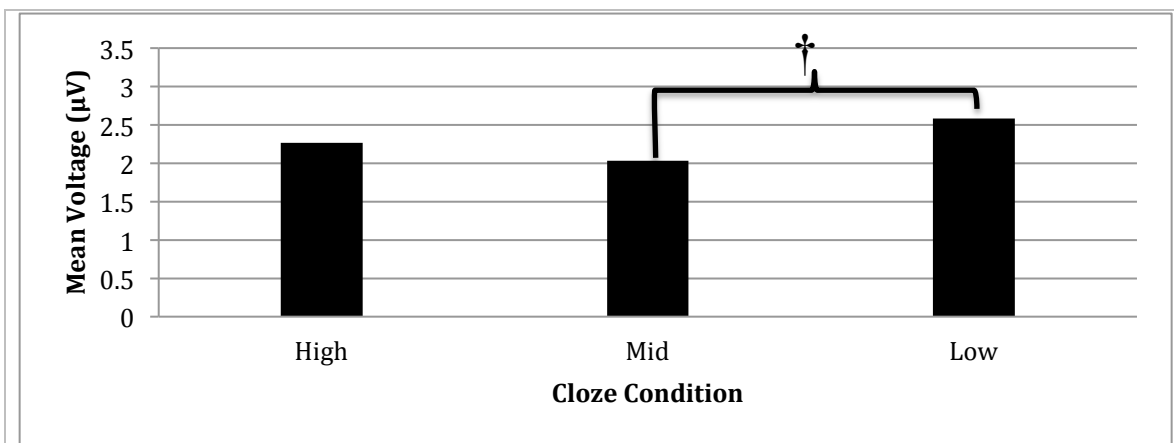


Figure 6. Mean voltage of cloze conditions within time window following Target-2 (140-340 ms).

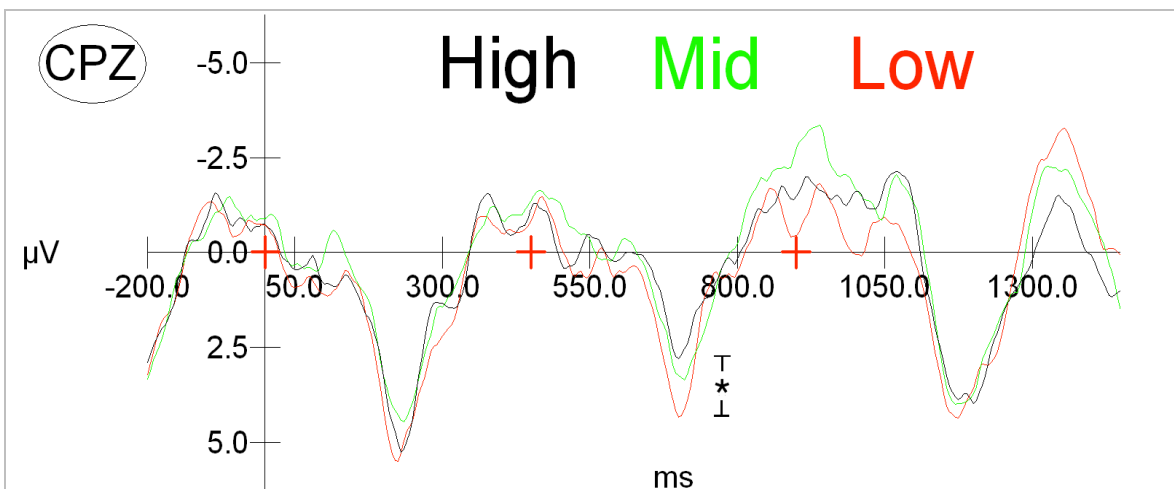


Figure 7. Pre-target EEG⁹ (see Footnote 9 for details on the proper interpretation of this graph of EEG).

⁸ The symbol † here denotes a finding with marginal statistical significance (i.e., $0.05 > p > 0.1$).

⁹ This snapshot of EEG activity is taken -200 ms to 1450 ms relative to the presentation Target-2. Thus, in addition to Target-2, the presentations of both Target-1 and the sentence-final target are

c. Target-1

A repeated measures ANOVA (again, collapsed across semantic anomaly and

Conditions compared	<i>t</i>	<i>df</i>	<i>p</i>
High-Mid [†]	-1.8785	16	0.07886
High-Low*	-02.2168	16	0.04147
Mid-Low	-1.4427	16	0.1684

Table 5: Paired *t*-test results for the Target-1 time window

well-formed control) was also conducted for this same component following the presentation of Target-1 (190-290 ms). Unlike the previous time window, a significant main effect of cloze probability was observed [$F(2, 32) = 3.9967, p < 0.05$]. More-

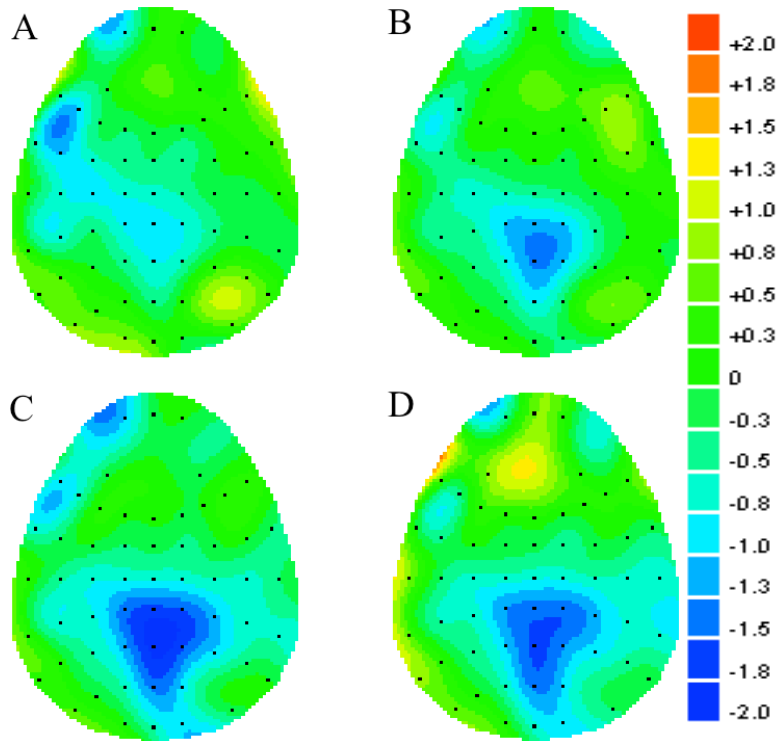


Figure 8. The scalp distribution of *t*-scores between the high and low cloze conditions plotted over time (A = 200-210 ms; B = 210-220 ms; C = 220-230 ms; D = 230-240 ms). The 200-240 ms range following the presentation of Target-1 was characterized by the greatest difference between the high and low conditions. Note that this difference is primarily localized within the pre-selected ROI (Figure 1).

over, paired *t*-tests of the mean voltages of each condition revealed that the difference between the high and low cloze conditions was significant and that the

difference between the high and mid cloze conditions was nearly significant, however no significance was found between the mid and low conditions. This information is summarized in Table 5 above. Figure 8 above represents the scalp distribution of the

also captured in this snapshot along with all the electrophysiological activity in between. The red crosshairs on the x-axis represent the points of presentation of Target-2 (0 ms), Target-1 (450 ms), and the target (900 ms), with a 450 ms SOA separating each.

significant difference over time between the high and low conditions (the only two conditions found to be statistically different enough to be worth visualizing) during the peak of the anticipatory component.

Note importantly the relative order of the means between the conditions (see Figures 9 and 10). Crucially, following the presentation of Target-1, the voltage of each

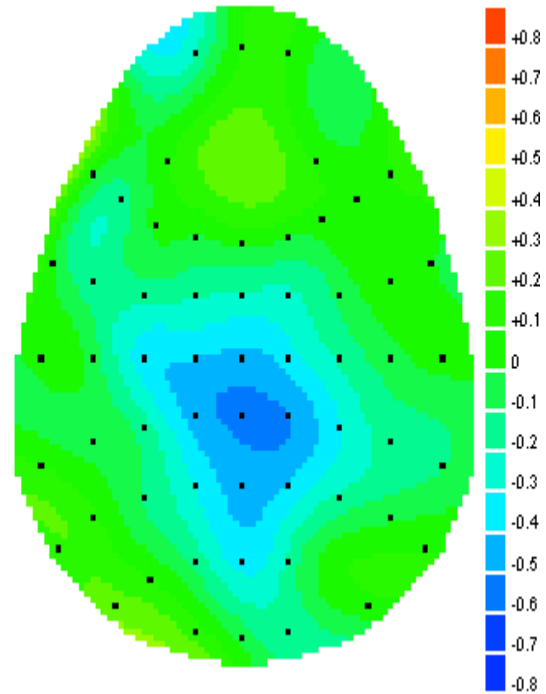


Figure 9. The scalp distribution of the correlation between post-Target-1 pre-target ERP effect size and cloze probability. After taking the mean voltage of each electrode within the 190-290 ms time window, the values of the three cloze conditions were input into a line graph at each electrode site. The slope of this line was then multiplied by r^2 to yield a measurement of the linear increase of the line with cloze condition. Though this topomap does not relate any information about statistical significance, regions closest to ± 1 thus represent the most significant ordering of effect size with respect to cloze probability.

condition within our time window appears to be ordered with respect to cloze probability – high cloze being the most negative, low cloze being the most positive, and mid cloze being between the two. To put it differently, the high cloze condition yielded the lowest amplitude

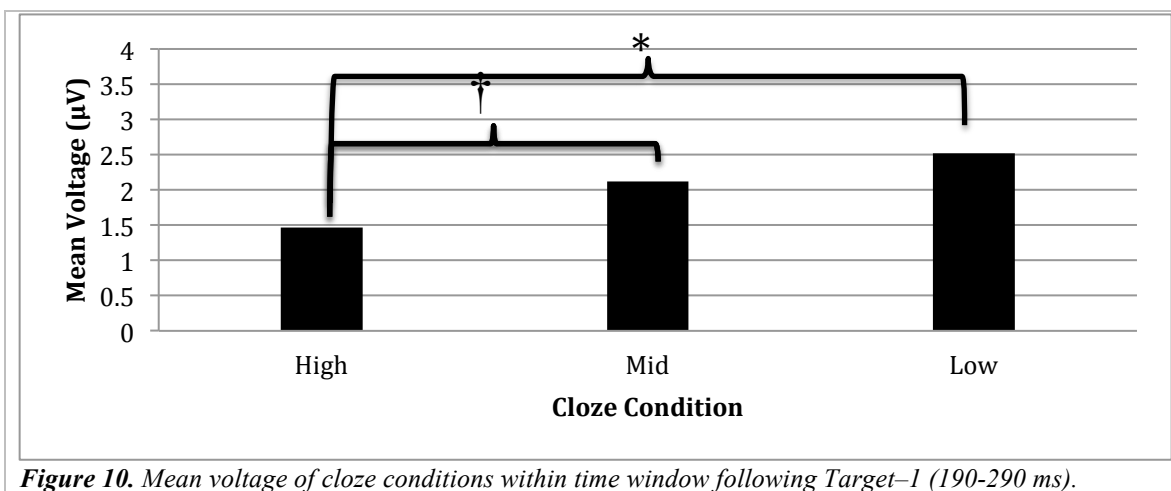


Figure 10. Mean voltage of cloze conditions within time window following Target-1 (190-290 ms).

ERP (as defined both by the absolute value of the peak and also the mean voltage of the wave throughout the 190-290 ms time window) following Target-1, while the low cloze condition yielded the highest amplitude ERP. Thus, within this time window, the effect size is inversely correlated with cloze probability. The paired t-tests reported in Table 5 confirm that of these three conditions, the difference between high and low is statistically significant, and the difference between high and mid is marginally significant. These mean voltage differences are thus likely a result of the manipulation of context constraint.

IV. Discussion

1. Direct pre-target evidence of anticipatory processing

a. Anticipatory activation and rapid semantic access

By manipulating context constraint, this study tested whether or not evidence of anticipatory processing could be directly observed in the pre-target region of EEG. The results suggest that context constraint does in fact affect pre-target processing, and does so by roughly 200 ms following the word immediately before the predicted element in a sentence. This finding is significant for two reasons: (1) As far as I am aware, no study has been previously conducted with the goal of identifying pre-target components directly (see Section I.5); (2) This provides evidence for early semantic access around roughly 200 ms post-stimulus onset, supporting the findings of some of the studies discussed in Sections I.1 and I.3.

Regarding (1), the anticipatory activation of element x in a discourse necessarily entails activating this entity prior to its actual utterance by the speaker, and consequently

the main processing effects of this prediction will occur prior to actually encountering the entity. As a result, in addition to the many extant studies that exclusively observe post-target ERP components to infer characteristics of pre-target neural activity, this study's primary focus on directly investigating the pre-target region allows it to make an important and original contribution towards understanding the precise mechanism behind the anticipatory processing of language; this in and of itself represents a significant first step in the direction classifying pre-target components that directly index aspects of predictive behavior important for language comprehension. While the design of this study has not revealed much in terms of the specific functional characteristics of the pre-target component identified here, it is important enough to show decisively that context constraint can influence the processing of elements downstream in the discourse, demonstrating that prediction on the basis of context constraint does indeed occur. This could segue directly into future research investigating this topic, as discussed more concretely in Section VI below.

Regarding reason (2) above, the findings of this study can be added to those introduced in Section I.3 which suggest that semantic processing may in fact be more rapid than previously thought to be the case. Specifically, at roughly 200 ms prior to encountering an item strongly predicted by a highly constraining context, semantic elements of this item may be anticipatorily activated, facilitating rapid integration into the discourse once this item is actually encountered. It is mechanisms such as this that may be the real reason behind why language processing, considering the tremendous computational complexity involved, is so rapid.

b. Temporal localization of the pre-target effect

The pre-target effect observed between the conditions was highly temporally localized. That is to say, in the entire pre-target region investigated here (i.e., Target–2 and Target–1), statistically significant differences between the cloze conditions were only observed to occur approximately 200 ms after the onset of Target–1, and this significance was lost again by about 300 ms. One may be tempted to ask in response to this distribution, if anticipatory activation supposedly represents the build up of linguistic constraint for a particular item, why is the distribution of this effect not more durable or widespread in time? Put differently, why do my data suggest that the context-dependent predictive mechanism seems active for only a 100 ms time window following the presentation of Target–1? Does it not seem more realistic to suppose that the resting activation level of a predicted item would remain elevated (and possibly even increase) from the beginning of the anticipatory processing period until the item is actually encountered, rather than dropping back to normal again after only 100 ms as these data seem to suggest? In fact, my original thinking was that the effect of increasingly constrained linguistic prediction would manifest itself in EEG in the form of cloze-correlated drift occurring throughout the presentation of multiple words. This is not what the data appear to show, and consequently the narrowly distributed temporal localization of the pre-target effect that actually was observed begs an explanation.

My account for this curious distribution relies on a feature of my experimental design, namely, the rapid SOA (i.e., presentation rate) of stimulus items. Each word in my stimulus sentences was on the screen for 200 ms with a 250 ms ISI (i.e., blank between words). This means that 450 ms after the presentation of one word, the following

word was displayed. As a result, post-stimulus onset activity that is approaching the 450 ms mark will be interrupted by the presentation of the subsequent word (or perhaps even by the anticipation that a word will be displayed, after subjects have become habituated to the presentation rate¹⁰). This could account, for example, for the discrepancy in the ordering of cloze conditions in the N400 region (300-500 ms) following Target-1 and that of the sentence-final target (see Figure 7 on page 26); namely, the post-target N400 region is characterized by a ranking of the voltages of each condition according to cloze probability (inversely mirroring the 190-290 ms window of Target-1), while the corresponding N400 region following Target-1 does not demonstrate any such ordering. It may have been the case that this orderly, cloze-determined pattern of activation after the target did in fact also characterize the post-Target-1 activity as well, but that after Target-1 this pattern was interrupted by the presentation (or rather the anticipation) of the target, and the otherwise clear anticipatory effect following Target-1 would have thus been muddled by the early prediction and subsequent sensory processing of the target. This would explain in Figure 7 why the context-constrained anticipatory effect appears localized to the 190-290 ms time window following Target-1. This concern could presumably be addressed by a future study that mimicked the one reported here but with a slower rate of presentation. This topic is explored further in Section VI.

¹⁰ The possibility that a subject's anticipation of an upcoming stimulus after becoming accustomed to the rate of presentation is an interesting thought. Moreover, it could have significant implications for the study of early components and predictive processing in that it may be difficult in the case of studies with regular (i.e., non-random) SOAs to tease apart EEG activity resulting from this anticipation and from activity resulting from the experimental manipulation. No studies were found that addressed this topic however, so currently the role of SOA habituation appears to be an open question and must be kept in mind when considering any interpretation of the components of anticipatory processing.

c. Post-Target-2 activity

Thus far all of the discussion in this section has dealt with the cloze-dependent EEG activity found to occur after Target-1. However, I have not yet addressed the issue regarding Target-2. The same component was observed to occur after both pre-target words, however the differences between the conditions were only statistically significant following Target-1. The question remains then, why was the same pattern of activation not found following Target-2? Part of this issue was implicitly addressed in Section IV.1.b above regarding the lack of a drift-like effect, namely that part of the explanation may lie in the rapid SOA of my experimental design. This does not fully explain the lack of significance still more than 200 ms in advance of Target-1 however, the time period around which the post-Target-1 activity is in fact significantly different between the conditions. Perhaps a more satisfying explanation may be that the anticipatory activation of a specific item or set of items in context may become maximal immediately before this item is predicted by the language processor to occur. It is at this point in an utterance when the semantic constraints not only call most strongly for the predicted item to occur, but may also maximally converge to this effect with the syntactic (McRae et al., 1998) and prosodic expectations (Grosjean & Hirt, 1996) as well, perhaps in addition to other sources of information not yet well understood. This integration of multiple sources of constraint immediately before the target word would account for why the effect is much more clearly manifested after Target-1 than Target-2. Moreover, this observation regarding the immediately pre-target integration of several kinds of predictive information may have considerable implications for the functional significance of this

component, however a detailed discussion of this possibility is beyond the scope of this study.

2. Directionality of the effect

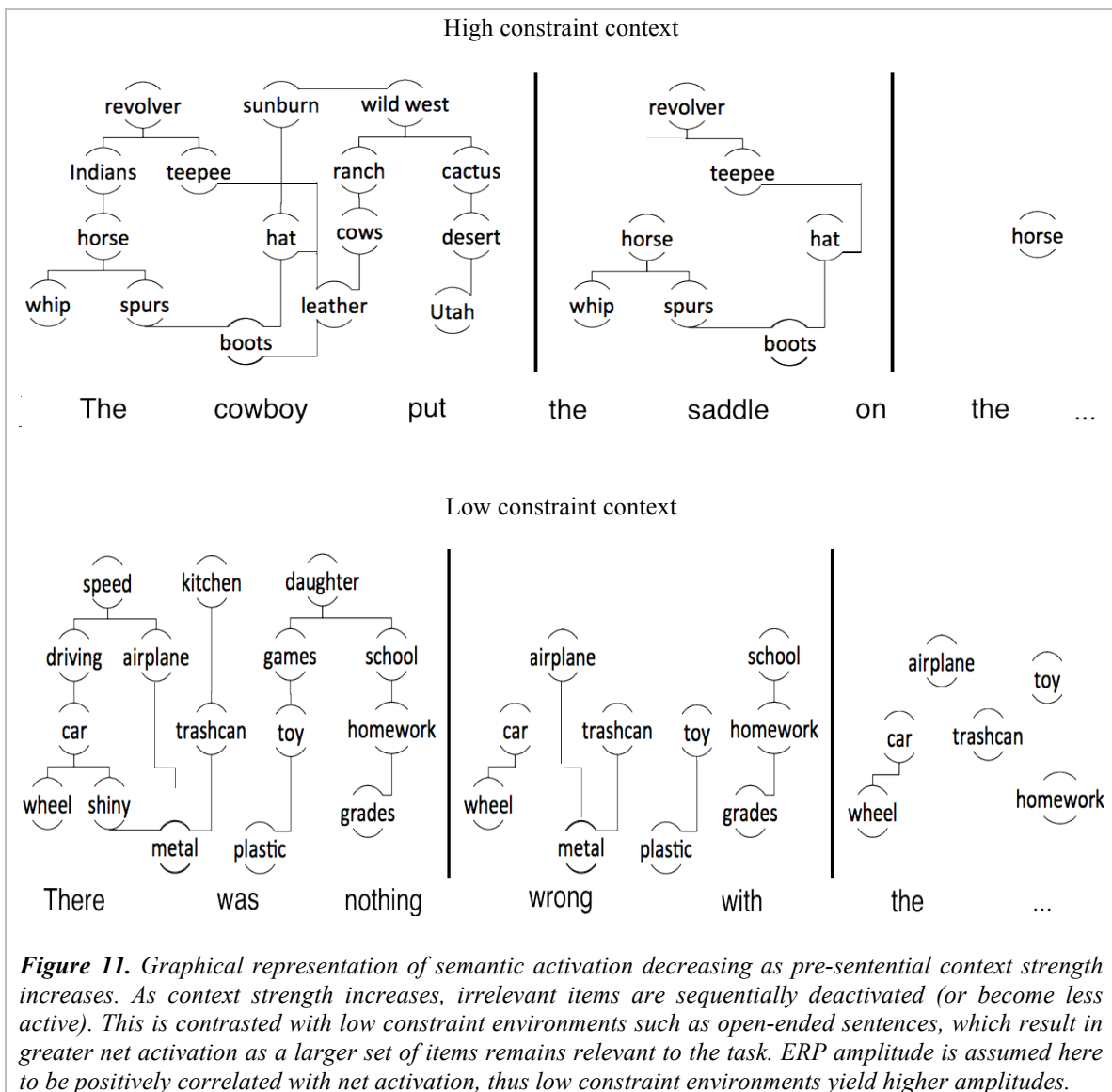
a. Ordering of cloze conditions in the pre-target region

In addition to the statistically significant difference between the mean voltages of the high and low cloze conditions within the effect following Target-1, it was also observed that these conditions were ordered with respect to cloze probability. Specifically, the high cloze condition yielded the lowest amplitude, followed by the mid and then the low cloze conditions. This numerical trend is quite significant for the interpretation of my results, in that it demonstrates that increasing levels of contextual constraint lead to lower levels of activation in the pre-target region (that is, that contextual constraint and anticipatory activation are inversely related). This finding accords with my hypotheses.

Increased amplitudes of an ERP component are commonly taken in the cognitive neuroscience literature to signify increased levels of activation, whether this activation represents either feature detection or difficulty of processing¹¹. Applying these insights to the findings of this study, because low cloze sentences were characterized by higher amplitudes of the anticipatory component in the pre-target region, it can thus be argued

¹¹ 'Feature detection' here refers to any neurocognitive system that activates in response to detecting certain preferred features of a stimulus. The amplitude of an ERP component that exemplifies feature detection logic will increase in the presence of one of these preferred features, and decrease in its absence. For example, the electrophysiological response to facial recognition is characterized by feature detection logic, because recognized faces beget strong levels of activation (Bentin et al., 1996). 'Difficulty of processing' logic is the reverse of this, namely, a pattern of electrical activation in the brain whereby increased amplitudes correspond to the difficulty of processing certain information. To put it trivially, feature detection systems respond more strongly to what they 'like' to see, while difficulty of processing systems respond more strongly to what they 'don't like' to see. In either case, higher ERP amplitudes are interpreted as higher levels of activation of these systems.

that weakly constraining sentential contexts require the listener to hold a larger number of possible completions active in the brain until the appropriate completion is actually encountered. Another way of stating this interpretation is that as the set of semantically appropriate completions decreases in size with increasing contextual constraint (i.e., as anticipatory activation becomes more specific), the total amount of neural network activation should also decrease as irrelevant nodes are sequentially pruned from the network (see Figure 11 above for a graphical representation of this process). This would



result in an overall lower level of activation for highly constraining contexts, which ultimately manifests itself in lower ERP amplitude. This is precisely what has been observed in the data reported here.

b. Relationship between N400 size and the pre-target effect

It was stated above several times that across the cloze conditions, the relationship between the amplitude of the pre-target anticipatory effect following Target-1 was observed to be inversely related to the size of the corresponding post-target N400 effect. This inverse relationship is in line with Hypothesis 2 as introduced in Section I.5.c. This is an interesting observation, but is not of much utility by itself. What it implies however is that, given a sufficiently specific model of the mechanisms behind anticipatory activation, and given a measurable N400 effect (i.e., the difference between a baseline and a semantically anomalous condition), one should be able to predict from the amplitude of the N400 effect alone the level of constraint that had characterized the preceding context. The converse should also be possible, in that given the cloze probability of a particular item within a sentence, one should be able to predict the size of the N400 effect that will result when the expectation of this item is violated. This relationship has potentially serious implications for the functional significance of the N400 effect, as discussed in the following section.

3. Reinterpretation of the N400 effect

When comparing N400 amplitude across both cloze and anomalous/control conditions, N400 amplitude for semantic anomaly appears following a visual inspection

to remain roughly constant between all three cloze conditions, while the amplitude of the well-formed control baseline varies markedly (see Figures 2-4). In fact, well-formed control amplitude increases as the semantic constraint of the pre-sentential context decreases, with the high cloze condition yielding the lowest amplitude, followed by the mid and then the low cloze conditions. Admittedly, no statistically significant interaction effects were observed between cloze condition and anomaly type, however the fact that the trend visually appears quite clearly to be maintained throughout all three conditions (that is, baseline N400 amplitude: high < mid < low) is curious, and begs the question of whether the lack of significance may be due to a flaw in the design of the experiment, or simply due to low amounts of data. These possibilities and others are discussed in Section V.1 below. Nevertheless, if this trend can be replicated with higher numbers of participants it may have considerable implications for the functional interpretation of the N400, and thus it will be preemptively included in the discussion here.

The N400 effect has traditionally been thought to index the integration of – or perhaps more accurately, the difficulty of integrating – new semantic information with the preceding context (Kutas & Federmeier, 2011). This view of the component has been reached by a wealth of studies for which semantically anomalous stimuli resulted in elevated N400 amplitudes relative to the baseline. Because anomalous items yielded relatively higher N400s, this pattern of activation was seen as suggesting that N400 size indexes the amount of difficulty for the language processor of integrating semantic information, with higher amplitudes reflecting greater processing difficulty. Consequently this understanding of the N400 effect has become the orthodoxy among contemporary cognitive neuroscientists.

However, the pattern of data reported here suggests a different interpretation of this effect. The fact that the level of the brain's response (i.e., N400 amplitude) to the presence of semantic anomaly does not seem to index the level of pre-sentential constraint or the subsequent degree of violation does not support the orthodox view of the N400. Rather, the fact that the baseline alone is what varied in response to changes in cloze probability suggests instead that the N400 effect may reflect not semantic integration or anomaly, but context strength, that is, the level of semantic constraint experienced by the language processor. This gets back to the point made in Section IV.2.b regarding the mutual predictability of cloze probability and N400 amplitude that is suggested by these data. Because one can approximately predict the N400 amplitude following well-formed utterances based on cloze probability, and conversely because one cannot predict N400 amplitude following semantically anomalous utterances, this implies that the N400 may more directly reflect the level of context constraint during normal (i.e., non-anomalous) language use than semantic processing difficulty.

V. Drawbacks and limitations

1. Low number of subjects

This simple experiment has provided a handful of suggestive insights into the nature of anticipatory language processing and the physiology of the N400 effect. However, like many studies, it also suffers from a number of drawbacks which merit candid discussion here. First, the number of subjects recruited for this study was unfortunately low. Although statistically significant results were found, 17 subjects is

quite a low number for an ERP experiment, as larger numbers of subjects are generally required in order for findings to be confidently generalizable over large populations. Moreover, some of the results that are approaching significance (e.g., the differences between the high and mid conditions and the mid and low conditions during the time window following Target-1) may have achieved it if the number of participants had been greater.

2. Insufficient control of Target-2 and Target-1

A second problem with the study is in the experimental design, namely, that the words in positions Target-2 and Target-1 were not controlled across sentences for variables known to affect online processing; such variables include grammatical class such as the noun/verb distinction (Federmeier et al., 2000; Hasting et al., 2008; Koenig & Lehmann, 1996) or the open/closed class distinction (Dikker et al., 2009; Dikker et al., 2010; Neville et al., 1992; Pulvermüller et al., 1994), semantic category such as animate/inanimate (Paczynski & Kuperberg, 2001; Phillipp et al., 2008; Szewczyk & Schriefers, 2011), as well as word length and frequency (Assadollahi & Pulvermüller, 1992; Osterhout et al., 1997). This is why cognitive neuroscientists and psycholinguists go to great lengths to rigorously control their stimuli prior to running an experiment. It is hoped that any potentially confounding effects of the linguistic inconsistency of the words in the Target-2 and Target-1 positions have been effectively washed out by the sheer variety of these words, and that what might otherwise be considered a confounding influence could be considered simply noise. However I cannot be certain of this assumption until further scientific investigation is conducted into the issue. This is

admittedly therefore a potentially significant flaw in my experimental design, however it could not be avoided here because, due to the time constraints of putting together this study, the stimuli sentences were largely borrowed from other studies (Bloom & Fischler, 1980; Hamberger et al., 1996; Block & Baldwin, 2010; Kim & Lai, In Press) which did not share the same experimental objectives.

3. Mean length between cloze conditions

A third potential drawback, intimately related in origin to the design flaw just described, is that the majority of the prefabricated stimulus sentence contexts may not have been sufficiently controlled for length between the cloze conditions. Table 1 summarizes the length means and ranges for each condition; high cloze sentences are on average the longest, with low cloze sentences being on average the shortest. Some readers of this report may be tempted to object in response that context constraint tends to increase with the length of a sentence. This particular argument should not be of major significance to my study, because although sentences in the high cloze condition are on average longer than those in the low cloze condition, this trend corresponds with the cloze conditions, and thus should not confound the conclusions drawn here. However, there may be other problems for my study associated with improperly controlled sentence length that are currently unclear, and this possibility must be kept in mind when pondering the results of this experiment.

4. Underspecification of findings

Finally, while not a drawback per se, this study is limited in terms of the specific conclusions it is able to draw regarding the precise nature of the anticipatory processing mechanism or of the information that is actually being anticipatorily activated during online language comprehension. That the pre-target cloze-dependent component tentatively identified may be involved in the integration of multiple sources of prediction (as discussed in Section IV.1.c) is a fascinating possibility, however this study alone is insufficient to make this determination. Because the issue addressed by this study appears to be truly unprecedented in the literature, I began with essentially the most basic question: Can we identify any direct evidence of anticipatory processing in the pre-target region, rather than relying exclusively on post-target activity? The answer to this question now appears to be a tentative yes, however at the moment the issue is inchoate and underspecified, and it will be up to future research to apply the findings reported here to ask narrower and more informative questions.

VI. Directions for future research

In response to both the findings and the limitations of the experiment reported in this paper, two directions for future research which build off of this study are readily apparent. First, an experiment could be conducted which tries to replicate these same findings but with a slower presentation rate, as suggested in Section IV.1.b. This design would help uncover more completely the physical form of the pre-target anticipatory effect reported here, and specifically it would address the question discussed in several

places above as to whether this effect truly is as temporally localized as these data suggest (i.e., 190-290 ms after the pre-target word only) or if it may actually be more appropriately characterized as drift over a longer period of time as was hypothesized. In this case, presumably the entire design of this study can be borrowed and recycled with a longer SOA, however one may wish to seriously consider the second and third drawbacks of my design discussed previously in Section V.

Second, because this study suggests that anticipatory activation can be directly observed in the pre-target EEG but is too premature to tell us anything about the actual details of this mechanism, a series of experiments needs to be designed to help uncover the specific details of this predictive mechanism. This could be done by rigorously controlling stimuli along various dimensions to determine more precisely which semantic features (or other kinds of features for that matter) are anticipatorily activated and in what time course. This would help us learn, for instance, whether there was only one anticipatory processing mechanism or several that operate along different linguistic dimensions (e.g., the pre-activation of syntactic vs. semantic vs. phonological structures), or to what degree predictive information about an item was integrated prior to actually encountering that item rather than after its presentation as has traditionally been the object of study. It would also be interesting to launch an investigation into the types of information structures that facilitate the performance of this anticipatory processor. That is to say, what features of language maximally facilitate or disrupt the normal functioning of this anticipatory processor, and what does this tell us about the way language is processed more generally?

VII. Conclusion

This study manipulated context constraint in order to uncover direct evidence of anticipatory processing in the pre-target EEG. Both hypotheses introduced in Section 1.5.c were confirmed by the data, namely (1) that N400 amplitude is directly proportional to context strength, and (2) that the amplitude of the pre-target ERP effect of anticipatory processing is inversely proportional the amplitude of the post-target N400. Beyond this, a number of other findings were discussed which are of significance to the study of anticipatory processing and to the interpretation of the N400 effect. Drawbacks of this study were also touched upon, as well as ways in which future research can build off of the results reported here.

I feel that the outcome of this experiment, while admittedly tentative and underspecified at present, contributes by pointing in useful directions for a heightened understanding of the ways in which linguistic structures are anticipatorily activated during online sentence comprehension. In particular, these results have begun the identification of roughly what the pre-target ERP components of anticipatory activation look like, where in the pre-target EEG researchers should concentrate their investigations to find the clearest components, and what the directionality of these components may imply (i.e., that more prediction leads to lower levels of pre-target activation).

The methodology of this experiment is unprecedented in that it appears to be the first to directly investigate the pre-target EEG directly for evidence of anticipatory processing, however such a focus is absolutely necessary in order to fully characterize the nature of the anticipatory processing mechanism. My findings alone are insufficient for a full characterization, however when applied in tandem with the many other studies that

have asked similar questions by utilizing only post-target neural activity, I feel that a more productive dialogue can begin to occur between these distinct yet mutually informative methodologies.

References

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11* (9), 1109-1116.
- Altmann, G. T. M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73* (3), 247-264.
- 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.
- Aruin, A. S. & Latash, M. L. (1995). The role of motor action in anticipatory postural adjustments studied with self-induced and externally triggered perturbations. *Experimental Brain Research*, *106* (2), 291-300.
- Assadollahi, R. & Pulvermüller, F. (2001). Neuromagnetic evidence for early access to cognitive representations. *Neuroreport*, *12* (2), 207-213.
- Assadollahi, R. & Pulvermüller, F. (2003). Early influences of word length and frequency: A group study using MEG. *Neuroreport*, *14* (8), 1183-1187.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Science*, *11* (7), 280-289.
- Ben Shalom, D. & Poeppel, D. (2008). Functional anatomic models of language: Assembling the pieces. *Neuroscientist*, *14* (1), 119-127.
- Bendixen, A., San Miguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83* (2), 120-131.
- Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*, 343-355.

- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8 (6), 551-565.
- Block, C. K. & Baldwin, C. L. (2010). Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. *Behavioral Research Methods*, 42 (3), 665-670.
- Bloom, P. A. & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, 8 (6), 631-642.
- Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *PLoS ONE*, 4, e5047.
- Deacon, D., Mehta, A., Tinsley, C., & Nousak, J. M. (1995). Variation in the latencies and amplitudes of N400 and NA as a function of semantic priming. *Psychophysiology*, 32 (6), 560-570.
- Dehaene, S. (1995). Electrophysiological evidence for category-specific word processing in the normal human brain. *Neuroreport*, 6 (16), 2153-2157.
- 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.
- Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, 110 (3), 293-321.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21 (5), 629-634.
- Duffy, S. A. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12 (2), 208-219.

- Federmeier, K. D. & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory & Language*, 41 (4), 469-495.
- Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain*, 123, 2552-2566.
- 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.
- Fowler, C. A. & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, 36 (2, 3), 171-195.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1 (3), 183-192.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science*, 6 (2), 78-84.
- Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31 (5), 862-877.
- Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, 22 (4), 477-492.
- Geschwind, N. (1967). The varieties of naming errors. *Cortex*, 3 (1), 97-112.
- Gibbon, F., Hardcastle, W., & Nicolaidis, K. (1993). Temporal and spatial aspects of lingual coarticulation in /kl/ sequences – A cross-linguistic investigation. *Language and Speech*, 36 (2-3), 261-277.

- Grosjean, F. & Hirt, C. (1996). Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. *Language and Cognitive Processes, 11* (1-2), 107-134.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural-network model of speech perception. *Psychological Review, 102* (3), 594-621.
- Hamberger, M. J., Friedman, D., & Rosen, J. (1996). Completion norms collected from younger and older adults for 198 sentence contexts. *Behavior Research Methods, Instruments, & Computers, 28* (1), 102-108.
- Hasting, A. S., Winkler, I., & Kotz, S. A. (2008). Early differential processing of verbs and nouns in the human brain as indexed by event-related brain potentials. *European Journal of Neuroscience, 27* (6), 1561-1565.
- Indefrey, P. & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92* (1-2), 101-144.
- Jerde, T. E., Soechting, J. F., & Flanders, M. (2003). Coarticulation in fluent fingerspelling. *Journal of Neuroscience, 23* (6), 2383-2393.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003a). 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. (2003b). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32* (1), 37-55.
- Kim, A. & Lai, V. (In Press). Rapid interactions between lexical access and word-form analysis during word recognition in context: evidence from ERPs. *Journal of Cognitive Neuroscience, 1-9*.
- Koenig, T. & Lehmann, D. (1996). Microstates in language-related brain potential maps show noun-verb differences. *Brain and Language, 53* (2), 169-182.

- Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207 (4427), 203-208.
- Kutas, M. & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307 (5947), 161-163.
- Kutas, M. & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4 (12), 463-470.
- Kutas, M. & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62. 621-647.
- Langacker, R. W. (1987). *Foundations of cognitive grammar – Vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.
- Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98 (1), 74-88.
- Lubker, J. & Gay, T. (1982). Anticipatory labial coarticulation – Experimental, biological, and linguistic variables. *Journal of the Acoustical Society of America*, 71 (2), 437-448.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge: MIT Press.
- 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.
- Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42 (4), 313-329.
- Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3 (2), 151-165.

- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2 (3), 244-258.
- Nieuwland, M. S. & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18 (7), 1098-1111.
- 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., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: word length and frequency predict the latency of an early negativity. *Biological Psychology*, 46 (2), 143-168.
- Otten, M. & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction of priming? *Discourse Processes*, 45 (6), 464-496.
- Paczynski, M. & Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing. *Language and Cognitive Processes*, 26 (9), 1402-1456.
- Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology*, 74 (3), 374-388.
- Phillipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language*, 105 (2), 112-133.
- Pratarelli, M. E. (1995). Modulation of semantic processing using word length and complexity: An ERP study. *International Journal of Psychophysiology*, 19 (3), 233-246.

- Price, C. J. (2000). "Functional imaging studies of aphasia." In: Mazziotta, J. C., Toga, A. W., & Frackowiak, R. S. J., eds. *Brain mapping: The disorders*. San Diego: Academic Press. P. 181-200.
- Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Electrocortical distinction of vocabulary types. *Electroencephalography and Clinical Neurophysiology*, 94 (5), 357-370.
- Pulvermüller, F., Assadollahi, R., & Elbert, T. (2001). Neuromagnetic evidence for early semantic access in word recognition. *European Journal of Neuroscience*, 13 (1), 201-205.
- Schwanenflugel, P. J. & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24 (2), 232-252.
- 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.
- Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, 23 (6), 695-703.
- St. George, M., Mannes, S., & Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience*, 9 (6), 776-787.
- Strange, W. & Bohn, O. S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *Journal of the Acoustical Society of America*, 104 (1), 488-504.
- Szewczyk, J. M. & Schriefers, H. (2011). Is animacy special? ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research*, 1368, 208-221.

- Tanji, J. & Evarts, E. V. (1976). Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *Journal of Neurophysiology*, 39 (5), 1062-1068.
- Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Journalism Quarterly*, 30 (4), 1953, 415-433.
- Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11 (6), 657-671.
- 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. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003a). Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, 346 (3), 165-168.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003b). 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 (3), 483-508.
- Wicha, N. Y. Y., Moreno, E. M., & 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*, 16 (7), 1272-1288.

Appendix: Experimental stimuli and filler sentences

High Cloze				
	Context	Control	Anomaly	Cloze
1	His breakfast came with eggs and three strips of	bacon	states	1.000
2	The vampire tried to bite the girl's neck to suck her	blood	law	1.000
3	The librarian searched the shelf for the	book	barrel	1.000
4	The patient was unconscious for too long and might have damage to his	brain	home	1.000
5	After the prisoner ate lunch he was led back to his	cell	air	1.000
6	At the end of dinner the customer paid the bill by writing a	check	car	1.000
7	The third grade teacher Mrs. Smith has twenty students in her	class	gift	1.000
8	The student wondered what time it was and looked to the back of the classroom to glance at the	clock	moves	1.000
9	Half-&-half is a mixture of milk and	cream	hotels	1.000
10	On his head the king wore a	crown	bottle	1.000
11	Many cancer researchers are trying to find the	cure	hat	1.000
12	After gathering his courage he asked the girl out on a	date	trout	1.000
13	In his speech Martin Luther King said he had a	dream	market	1.000
14	The boy scouts forgot to bring matches so they rubbed sticks together to start a	fire	nurse	1.000
15	Apples and oranges are two different kinds of	fruit	news	1.000
16	His cute niece said goodbye with a wave of her	hand	pillow	1.000
17	Using a stethoscope the doctor can hear the beating of his	heart	sink	1.000
18	The cowboy put a saddle on the	horse	percent	1.000
19	He can't bend his leg because he injured his	knee	kind	1.000
20	Nine months into the pregnancy the mother went into	labor	man	1.000
21	When the man walked into the dark room he turned on the	light	disease	1.000
22	In order to buy the car he went to the bank to get a	loan	class	1.000
23	At first sight he fell in	love	olive	1.000
24	To catch up on the day's events Mary's parents watched the	news	gods	1.000

25	During the removal of the wisdom tooth the dentist gave him a shot to ease the	pain	book	1.000
26	He wrote down his thoughts on a piece of	paper	trouble	1.000
27	His mom told him to eat all the food on his	plate	ceiling	1.000
28	The shop needed to sell the item so they reduced the	price	ropes	1.000
29	He proposed marriage with the diamond engagement	ring	face	1.000
30	Racists judge people by the color of their	skin	oil	1.000
31	After the blizzard all the cars were buried in	snow	meat	1.000
32	The main course comes with either a salad or a bowl of	soup	trash	1.000
33	At the end of the play the actors all came back on	stage	prison	1.000
34	After the other boy told his account of what happened they let her tell her side of the	story	tub	1.000
35	She sipped the iced tea through a	straw	test	1.000
36	After visiting the dentist he promised to always brush his	teeth	songs	1.000
37	While visiting the college the students and parents were taken on a	tour	seat	1.000
38	When John's parents thought he was lying they asked him to tell the	truth	bar	1.000
39	He opened an unknown application that infected his computer with a	virus	cloud	1.000
40	Before going into the bathtub she tested the temperature by dipping her toes into the	water	murder	1.000
41	The guy had a weird look on his	face	ink	0.981
42	The teenager's parents were not around last weekend so he threw a	party	line	0.981
43	The kid built a sand castle on the	beach	money	0.980
44	The prosecutor provided evidence that convinced the jury beyond the shadow of a	doubt	head	0.980
45	By crossing the finish line the car driver finished the	race	charge	0.979
46	The outfielder dove and caught the	ball	car	0.977
47	As to where he left his keys he hasn't got a	clue	bush	0.977
48	The janitor whistled while he mopped the	floor	wall	0.977
49	The mother bird laid eggs in the	nest	shelf	0.977
50	The old porch looked completely new with a fresh coat of	paint	pork	0.977
51	He didn't have time for an entire meal so he had crackers and cheese for a	snack	port	0.977
52	The backpacker finally found a camp site and set up the	tent	search	0.977

53	At the wedding reception the best man raised his glass for a	toast	screen	0.977
54	After the funeral they laid him to	rest	plane	0.977
55	He grew up in the country but would like to live in the	city	card	0.976
56	The basketball players were told not to stay up late the night before the	game	step	0.976
57	After hearing the ridiculous question Janet smiled and shook her	head	plane	0.976
58	She gambled online and lost a lot of	money	glasses	0.976
59	The crying baby kept her parents up all	night	mouse	0.976
60	She hung the picture on the	wall	risk	0.976

Mid Cloze				
	Context	Control	Anomaly	Cloze
61	Forks and knives are kinds of	utensils	days	0.625
62	The doctor solemnly told Joshua that he had	cancer	laundry	0.625
63	Cakes are the most delicious kind of	dessert	skill	0.625
64	Jack stealthily snuck out of the room without a	sound	duty	0.610
65	The boy and his girlfriend went on a date to a	movie	printer	0.600
66	He was way off of the main	road	golf	0.600
67	The fish jumped right out of the water and into my	boat	work	0.600
68	A large stone blocked the entrance to the	cave	idea	0.600
69	After seeing the movie she felt	sad	federal	0.600
70	The doorman hailed Mrs. Jones a	cab	court	0.600
71	There were only three days until his	birthday	phone	0.600
72	People at the next table came over and asked if someone was sitting in that	chair	cancer	0.579
73	He wondered whether it was coincidence or	fate	dessert	0.567
74	Of all her favorite people Betsy liked Andy the	best	sound	0.567
75	The swimmer dove into the	pool	mustard	0.567
76	The sweater was not worn much because it collected	lint	pizza	0.567
77	The crazy girl destroyed her boyfriend's	car	marathon	0.567
78	The snowboarder jumped from the helicopter and began his	descent	cheese	0.550
79	The sandwich wasn't very good without a slice of	cheese	gum	0.550
80	He put his feet up on the	table	script	0.550
81	The child went even higher on the	swing	towel	0.550

82	By betting all her money on the new company she took a	risk	fury	0.547
83	He bought a new car and took the family for a	ride	plan	0.542
84	The couple was seated at the table by the	window	dirt	0.533
85	The hungry owl searched the field for its	prey	lamp	0.533
86	She did not have the right key for the	door	mouth	0.533
87	The boy asked his teacher for extra	credit	glass	0.533
88	Software is rigorously tested before it is	released	eaten	0.533
89	She took a picture with her	camera	guard	0.533
90	The recipe calls for two tablespoons of	sugar	sand	0.525
91	Eating lots of ice cream gave me a	headache	store	0.525
92	Cleopatra was a beautiful	woman	porch	0.525
93	Jumping rope can be a fun	activity	bread	0.525
94	The cruiseship didn't foresee the change of weather near the tropical island and was hit by a	storm	child	0.516
95	He became deeply absorbed in the novel after reading the first	chapter	judge	0.500
96	After parking behind the condo the movers unloaded the furniture out of the	van	tree	0.500
97	Renewed fighting has broken out near the	border	fingers	0.500
98	The squirrel scurried away from the	tree	bone	0.500
99	Since the year 2000 LA has seen a steady rise in	crime	food	0.500
100	The best remedy for the common cold is	rest	state	0.500
101	Marissa was still tired so she went back to	sleep	heaven	0.500
102	Lois is taller than most	women	years	0.500
103	The Jones children wanted a house with a large	yard	pride	0.500
104	When the employee was hired he began his new	job	fact	0.500
105	Sarah spent all summer lounging at the	pool	horse	0.500
106	For thousands of years people have used animals for	food	snow	0.500
107	The cabin was made of	logs	dishes	0.500
108	Bob was unhappy because his new raise was so	small	purple	0.500
109	The prince prepared for his coronation by first putting on his	crown	habit	0.475
110	Before it got dark the cowboys herded those ten cows back into the	barn	toast	0.472
111	They were startled by the sudden	noise	barn	0.467
112	Many books have pictures and	words	butter	0.467
113	She was late so she hoped the bus would come	soon	especially	0.467
114	Not even the cast liked the	play	health	0.467
115	The river ended three hundred miles from its	source	room	0.467
116	The taxi driver waited impatiently for his	fare	skirt	0.467

117	The friends skated across the frozen	pond	drug	0.467
118	The editor allowed the final draft to go to	print	rock	0.467
119	The cheerleaders were part of the popular	crowd	lake	0.467
120	In the play, Romeo and Juliet shared a passionate	kiss	beer	0.467

Low Cloze				
	Context	Control	Anomaly	Cloze
121	The sail came loose, so they tightened the	ropes	branch	0.250
122	Hiding in the tall grass, the lion spotted a	gazelle	president	0.250
123	When making a stir fry, the most important ingredient is	oil	news	0.250
124	The most terrifying animal in the world is the	shark	bridge	0.250
125	He used a screw driver to open the	box	world	0.233
126	The state prison is where the killer was	incarcerated	looked	0.233
127	Disney is famous for its	cartoons	speed	0.233
128	The gym teacher knew the rope would be a good	challenge	spoon	0.233
129	People often express their emotions through	song	cups	0.233
130	The husband saved his wife a	seat	land	0.233
131	This song was written for	her	beans	0.233
132	The airplane went into a	cloud	noise	0.225
133	Afghanistan has one of the highest rates of	murder	light	0.225
134	Police found the body in a	ditch	plate	0.225
135	The little girl's favorite food was	pizza	shoes	0.225
136	The internet has made it much easier for people to find	information	cabbage	0.225
137	Ample food was made for the	party	fate	0.225
138	David's car is broken so he has to go buy a	bike	mile	0.225
139	The director was committed to his artistic ideals and refused to change the	script	friends	0.224
140	One of my favorite hobbies is	reading	breathing	0.200
141	Credit cards are often used to buy	stuff	whales	0.200
142	The elderly sometimes lose their	memory	puzzles	0.200
143	This weekend I really want to go to the	movies	weather	0.200
144	Dinner needed serious rethinking after Shelly overcooked her	chicken	machine	0.200
145	Family reunions are important functions for	families	bushes	0.200
146	The suspect was last seen wearing a red	jacket	picture	0.200
147	I saw a raccoon last night in the	garbage	job	0.200
148	The ruby was so big it looked like a	strawberry	shelf	0.200

149	Ants were seen all over the	ground	direction	0.200
150	Boarding the ship, the old woman said goodbye to her	family	size	0.200
151	Kim wanted to give the workout class a	try	language	0.200
152	The main source of pollution today is	cars	salmon	0.200
153	Ordinary people could not afford to buy	gold	death	0.200
154	After failing he realized he needed a new	plan	beach	0.200
155	Nobody knows who made the first	wheel	potato	0.200
156	Let's go sit by the	fire	pepper	0.200
157	The boy stole a pencil because of a	dare	hair	0.200
158	The bouquet of flowers was pretty and smelled	sweet	popular	0.200
159	Mr. Hall hoped that the appointment would be	short	hungry	0.200
160	Scotland is famous for its	bagpipes	fiber	0.175
161	For an athlete Ted is rather	small	united	0.167
162	As time goes on computers are becoming more	advanced	blue	0.167
163	There are many different kinds of	people	internet	0.167
164	The boy refused to finish his dinner because he didn't like the	vegetables	nails	0.250
165	Countries like Brazil have many	people	floors	0.150
166	The Anderson's had never visited that	museum	math	0.150
167	Jazz music is often characterized by	saxophones	candy	0.133
168	Combing my hair makes me look like a	girl	race	0.133
169	The stars at night are very bright in	Texas	murder	0.133
170	Most people prefer not to eat	worms	tea	0.133
171	People disagree on how to define	love	carpet	0.133
172	The worst place to find oneself is in	trouble	strategy	0.250
173	Every culture has its own	ways	summer	0.133
174	There was nothing wrong with the	food	smoke	0.125
175	There is something wonderful about the	sky	curve	0.125
176	Their money was divided by the	banker	medicine	0.100
177	In New York there are two main	bridges	vacations	0.100
178	Jenny wanted him to bring a	blanket	school	0.100
179	I have never been to	Alaska	razor	0.075
180	This chapter is about the life of	Jesus	point	0.067

Fillers			
Filler sentence			
Comprehension question			Response
181	The dentist recommends brushing your teeth twice a day.		
	It's healthy to brush more than once a day.		True
182	Rob was in a hurry and kept watching the clock.		

	Rob was almost running late.	True
183	He crept into the room without a sound.	
	He was very stealthy.	True
184	Since the student missed the class he asked his doctor to write him a note.	
	The student was sick.	True
185	By beating the bully up he put him in his place.	
	Justice was served.	True
186	To pay for tuition she took out two student loans.	
	She didn't have enough money for tuition.	True
187	She didn't have her watch so she asked for the time.	
	She needed to know the time.	True
188	The man was severely injured and the paramedics struggled to save his life.	
	The man was in an accident.	True
189	To hang the picture Ted needed a hammer and nail.	
	Ted wanted the picture on his wall.	True
190	After the argument Ann went to her room and slammed the door.	
	Ann was unhappy with something.	True
191	On his 30th wedding anniversary James bought a necklace for his wife.	
	James had been married for three decades.	True
192	Upon seeing the enemy targets the general ordered the jet to drop the bomb.	
	The jet was on a combat mission.	True
193	After working the entire summer the young man had learned the tricks of the trade.	
	The young man gained valuable experience.	True
194	The witness gave his testimony during the trial.	
	The witness went to the stand.	True
195	Father carved the turkey with a knife.	
	It was Thanksgiving.	True
196	She graduated at the top of her class.	
	She is very intelligent and hard-working.	True
197	Expecting Jeff's call she waited for the phone to ring.	
	She was sitting by the phone.	True
198	The genie promised the man he would grant one wish.	
	The man had found a magic lamp.	True
199	The teacher wrote the problem on the board.	
	School was in session.	True
200	At night the old woman locked the door.	
	The woman lived in an unsafe neighborhood.	True
201	In the quiet movie theater, Kim's phone rang.	
	Kim didn't plan ahead.	True
202	It was dark in the room so she turned on the light.	
	It was nighttime.	True
203	When the alarm rang the firefighter slid down the pole.	
	The firefighter was at the station.	True
204	He cashed his new paycheck at the bank.	

	He has a job.	True
205	Cold weather outside meant it was time to turn on the heat.	
	It was autumn.	True
206	At first the woman refused, but she changed her mind.	
	The woman consented.	True
207	Because there was lightning she could not go to the pool to swim.	
	It was stormy outside.	True
208	The business man was good at negotiating and closed the deal.	
	The business man is persuasive.	True
209	Josh needed a belt to hold up his pants.	
	Josh is thin.	True
210	After playing in the sun all day, his face was badly burnt.	
	He was not wearing sunscreen.	True
211	Sally knew Julia was frightened because she looked scared.	
	Sally was frightened.	False
212	The woman took Tylenol to relieve her severe back pain.	
	The woman had injured her shoulder.	False
213	He lit the charcoal and then put the burgers and hot dogs on the grill.	
	The man was using a gas grill.	False
214	The salesman helped Sherry get a great deal.	
	The salesman was untrustworthy.	False
215	They wanted their parents to come home.	
	Their parents were asleep upstairs.	False
216	The laundry detergent failed to get out the red stain.	
	The grass left a permanent stain.	False
217	Glasses greatly improved Mrs. Gibb's sight.	
	Mrs. Gibb had decided to switch to contacts.	False
218	The police arrested the suspect who committed the crime.	
	The police got the wrong man.	False
219	He accidentally burned his tongue with the hot tea and lost the sense of taste.	
	He was drinking coffee.	False
220	Dan was asked to be the new coach of the team.	
	Dan was about to be replaced as coach.	False
221	They paid for their meals but forgot to leave a tip.	
	They left without paying.	False
222	The waiter in the Italian restaurant suggested a special bottle of wine.	
	The restaurant served Chinese food.	False
223	The boy didn't walk until his first birthday when he finally took a step.	
	The boy is eight months old.	False
224	He dipped his cookie in a glass of milk.	
	He is allergic to dairy.	False
225	The blocked chimney filled the cabin with smoke.	
	The cabin was on fire.	False
226	Without her sunglasses the sun hurt Erika's eyes.	
	It was a rainy day.	False

227	The limping horse was obviously in much pain.	
	The horse had injured its neck.	False
228	He mailed the letter without a stamp.	
	The letter will make it to its destination.	False
229	She wore a colorful scarf around her neck.	
	The scarf was black and white.	False
230	The cheap pen ran quickly out of ink.	
	The person needed to sharpen their pencil.	False
231	In the shower he washed his skin with soap.	
	He took a bath.	False
232	She lied about losing her report card to hide her bad grades.	
	She is a good student.	False
233	For his date Tom bought a long stemmed rose.	
	Tom bought a beautiful bouquet for his date.	False
234	Her job was easy most of the time.	
	The woman is a neurosurgeon.	False
235	Jenny lit the candles on the birthday cake.	
	Jenny was the birthday girl.	False
236	Although Keith bowled well he did not have the highest score.	
	Keith bowled a perfect game.	False
237	To promote their album the band went on tour.	
	The band had recently stopped making music.	False
238	Bill jumped in the lake and made a big splash.	
	Bill didn't know how to swim.	False
239	I could not remember his name.	
	I had never met him before.	False
240	Wanting color in the room, he bought a can of paint.	
	He planned to paint the walls off-white.	False