WHICH WAY TO INTEGRATION? EXAMINING DIRECTIONAL ASSOCIATION IN WORD-TO-TEXT INTEGRATION USING ERPS

by

Joseph Z. Stafura

B.A., Purchase College, State University of New York, 2007

Submitted to the Graduate Faculty of

Kenneth P. Dietrich of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

Kenneth P. Dietrich School of Arts and Sciences

This thesis was presented

by

Joseph Z. Stafura

It was defended on

March 8, 2013

and approved by

Natasha Tokowicz, Associate Professor, Department of Psychology

Tessa Warren, Associate Professor, Department of Psychology

Thesis Advisor: Charles Perfetti, Distinguished University Professor, Department of Psychology

WHICH WAY TO INTEGRATION? EXAMINING DIRECTIONAL ASSOCIATION IN WORD-TO-TEXT INTEGRATION USING ERPS

Joseph Z. Stafura, M.S.

University of Pittsburgh, 2013

Skilled reading proceeds in a largely incremental manner, with readers attempting to integrate linguistic information from each word as it is encountered. The degree to which prospective and retrospective integration processes are functional in driving incrementally remains an open area of inquiry. In this study, event-related potentials (ERPs) were used to explore on-line lexico-semantic integration in conditions in which prospective and retrospective processing was more or less likely to occur, through the manipulation of the direction of lexical association between word in isolation and embedded in two-sentence texts. The N400 ERP component, an index of lexico-semantic processing, was examined across forward and backward association conditions. In both a word relatedness judgment (RJ) task and text comprehension (TC) task, reduced N400 amplitudes were seen over central scalp electrodes in conditions in which word pairs were either forward associated or backward associated, relative to conditions in which word pairs were unrelated (RJ) or lacking one word of the pair (TC). Additionally, a reduced negativity was found for forward associated pairs over right parietal electrodes in RJ, and an increased positivity was found for the backward associated condition over left parietal electrodes in TC. The evidence from central electrodes suggests that retrospective integration processes, and not simply prospective expectancy processes, modulate the N400 in incremental text processing. Additionally, the results suggest an enhanced role for expectancy in modulating ERPs at right parietal sites, and, potentially, an engagement of memory resonance processes in text processing over left parietal sites.

Keywords: text integration, lexical association, relatedness judgments, N400, ERPs

TABLE OF CONTENTS

PREFACE X

1.0 INTRODUCTION 1

1.1 ON-LINE INTEGRATION IN LANGUAGE PROCESSING 2

1.2 SITUATION MODEL UPDATING 5

1.3 LEXICAL ASSOCIATION 8

2.0 CURRENT STUDY 13

2.1 HYPOTHESES 14

2.2 METHODS 16

2.2.1 Participants 16

2.2.2 Materials 17

2.2.3 Design and Procedure 20

2.2.4 Apparatus and ERP Recordings 24

3.0 RESULTS 26

3.1 DESCRIPTIVE DATA 26

3.2 TEXT COMPREHENSION 28

3.3 RELATEDNESS JUDGMENTS 32

3.3.1 Behavioral 32

3.3.2 ERP 33

3.4 INDIVIDUAL DIFFERENCE CORRELATIONS 36
3.4.1 Behavioral 36
3.4.2 ERP 38
4.0 DISCUSSION 40
BIBLIOGRAPHY 47

LIST OF TABLES

Table 1. Sample Passages from Yang, Perfetti, & Schmalhofer, 2007. 3

Table 2. Sample Passages for Each Experimental Condition 19

Table 3. Participant Descriptive Information.. 27

Table 4. Analysis of Variance (ANOVA) on Text Comprehension Task Mean Amplitude from 300-500ms at Cz, P3, and P4 Electrode Cluster Sites 31

Table 5. Relatedness Judgment Mean (Standard Deviation) Accuracy and Reaction Times....... 32

Table 6. Analysis of Variance (ANOVA) on Relatedness Judgment Task Mean Amplitude from 300-500ms at Cz, P4, P3 Electrode Cluster Sites 35

Table 7. Relatedness Judgment Reaction Time Data and Individual Difference Correlations 37

Table 8. Event-Related Potential (ERP) Mean Amplitude from 300-500ms and Individual Difference Correlations... 38

LIST OF FIGURES

Figure 1. Simple schematic of situation model updating. 7

Figure 2. Example sentence comprehension trial. 22

Figure 3. Waveforms and amplitude charts for the text comprehension task ERP data. 29

Figure 4. Waveforms and amplitude charts for the relatedness judgment task ERP data. 33

PREFACE

I have a great number of people to thank for support and guidance in pursuit of carrying out this project. First, my Advisor, Chuck Perfetti, has been an unending source of ideas, critiques, and support. Additionally, Tessa Warren and Natasha Tokowicz were of great help in the development and analysis of this project. A number of colleagues have been of invaluable support in a variety of ways, an abridged list follows (in no particular order): Suzanne Adlof, Laura Halderman, Ben Rickles, Li-Yun (Wendy) Chang, Alison Tseng, Jen Dandy, Fan Cao, and Jon-Michel Seman. The Learning Research and Development Center has been a great home for the past few years, and I've benefited from interactions with dozens of graduate students, faculty, and staff. Finally, my family is my greatest source of support, and always has been; thanks Dad, Mom, Dan, Penny, and Joce.

1.0 INTRODUCTION

For over three decades, psychological and psycholinguistic research has converged on the view that readers *incrementally* integrate information into their on-going mental representations of text (Just & Carpenter, 1980). That is, readers process words, to the extent possible, as they are encountered. Evidence for the immediate influence of words on comprehension processes has been found in behavioral responses in on-line listening (Tyler & Marslen-Wilson, 1977) and self-paced reading tasks (Boland, Tanenhaus, Garnsey, & Carlson, 1995), in eye-movements in visual world (Altmann & Kamide, 1999) and text reading tasks (Rayner & Clifton, 2009), and in even-related potentials (ERP) in passive reading tasks (Kutas, Van Petten, & Besson, 1988). Thus, at a number of levels (e.g. syntactic, semantic), linguistic features are used on-line by readers in the construction and updating of their understandings of text. Within this incremental view of reading, questions remain as to what processes are functional during on-line integration, and more specific to the focus of this study, which processes are functional during the integration of the semantic features of a word with its context? In this study, we used ERPs to contrast conditions in which prospective or retrospective integrative processes were more or less likely to occur, through the examination of scalp voltage potentials during a time-window known to reflect semantic processing.

1.1 ON-LINE INTEGRATION IN LANGUAGE PROCESSING

On one view of word-to-text integration during on-line incremental processing, readers use information in the text itself, as well as their background knowledge, to develop and constrain expectations of upcoming words. Upon encountering each word, then, more or less lexico-semantic processing is needed to fit the words into the context, given the match with prospectively developed expectations. Another view of integration is that, upon encountering and processing a given word, retrospective processes are engaged through which readers add the words semantic features to their representation of the previous text. These differential views of integration– prospective *vs* retrospective - are not likely to be mutually exclusive. Indeed, given the amount and structure of information in a text, it is reasonable to assume that skilled comprehenders utilize both types of processes when reading, with some texts lending themselves to greater or lesser amounts of prospection and retrospection.

A recent example of on-line word-to-text integration was shown in an ERP experiment carried out by Yang, Perfetti, and Schmalhofer (2007). In ERP studies, the continuous electroencephalogram (EEG) is recorded from electrodes placed at a number of locations on participants' scalps. Segments of the EEG, time-locked to events of interest (e.g., critical words), are averaged across trials. Averaging helps to cancel out random fluctuations of electrical activity unrelated to experimental stimuli, increasing the signal-to-noise ratio of the data. Additionally, the averaged trials are baseline corrected by subtracting mean activity occurring during a period prior to the event of interest, further attenuating EEG activity unrelated to the experimental manipulation. Finally, inferential statistics are applied to the ERP waveform data, generally comparing amplitudes or latencies of voltage deflections across experimental conditions.

Yang et al. (2007) utilized ERPs to examine the word-to-text integration processes of skilled comprehenders. The comprehenders read two-sentence passages, with ERP measurements taken during reading of the first content word of the second sentence (i.e., the *critical word*). The experimental manipulation was the critical word's referential availability in the first sentence. For example, in the *explicit* condition the critical word repeated a word in the first sentence (with occasional morphological variation), and in the *paraphrase* condition the critical word was conceptually related to a word or phrase in the first sentence. Example stimuli from Yang et al. (2007) are shown in Table 1.

Table 1. Sample Passages from Yang, Perfetti, & Schmalhofer, 2007.

Integration Type	Sample Passage
Explicit	After being dropped from the plane, the bomb hit the ground and exploded**.** The **explosion** was quickly reported to the commander.
Paraphrase	After being dropped from the plane, the bomb hit the ground and blew up**.** The **explosion** was quickly reported to the commander.
Baseline	Once the bomb was stored safely on the ground, the plane dropped off its passengers and left**.** The **explosion** was quickly reported to the commander.

Note. The critical word (explosion) is in underlined and bold at the beginning of the second sentence. The antecedent words in each condition are underlined.

The key finding from Yang et al. (2007) for this study was that in the paraphrase condition comprehenders' on-line processing of the critical word differed compared to a baseline condition in which the critical word had no available antecedent. This *paraphrase effect* was indexed by a reduction in amplitude of the N400 component elicited by the critical word in the paraphrase condition relative to that elicited by the critical word in the baseline condition. The N400 component is a negative deflection of the ERP waveform, maximal over centro-parietal electrodes and peaking in amplitude around 400 ms after exposure to any potentially meaningful stimulus. This component is modulated by the semantic fit between the currently processed stimulus and its context, with smaller amplitude waves reflecting a better fit (Kutas & Hillyard, 1980; Kutas & Federmeier, 2011). In the context of the Yang et al. (2007) study, then, critical words had a better lexico-semantic fit with the context when they were preceded by a conceptually-related word or phrase in the first sentence. Whether this ease of processing was a result of expectancy driven by the preceding context (Federmeier & Kutas, 1999), or of retrospective processes engaged during the processing of the critical word (Brown & Hagoort, 1993), is an open question.

1.2 SITUATION MODEL UPDATING

Before describing our manipulations in detail, it may be helpful to conceptualize our views of word-to-text processing through the lens of *mental models*, or, more specifically for reading, *situation models* (Johnson-Laird, 1981, 1983; Van Dijk & Kintsch, 1983). A situation model is a comprehender's representation of the meaning of a text or discourse, or of the situation that the text or discourse describes, rather than a verbatim representation of the surface-level features used to express this meaning. Research examining a range of dimensions represented in text - including time, space, causation, and protagonist intentionality - has revealed that comprehenders go beyond surface structure in the construction and updating of situation models by utilizing textual and knowledge-based referential and inferential processes (Graesser, Singer, & Trabasso, 1994; Zwaan & Radvansky, 1998). These processes influence readers' memory for aspects of a text, and help explain the functioning of situation models in enabling readers to integrate successive linguistic items across sentences and paragraphs in the face of semantic and referential in- or under-determinacy (Johnson-Laird, 1981).

In figure 1, we illustrate a simple situation model framework for the example from the Yang et al. (2007) study (Table 1). Upon reading the first sentence, the reader has an event structure in their situation model representing a bomb 'blowing up'. As the reader begins the second sentence, the new 'explosion' event can be linked to the co-referential event in the first sentence, leading to a new, integrated event structure of a bomb blowing up and an explosion occurring. That the antecedent event was already in the readers' mental model results in a facilitated updating process, relative to contexts lacking such co-referential events. At this point, our view of situation model updating is potentially consistent with both prospective and retrospective views on integration. First, it is plausible that the context of the first sentence enabled the readers to develop semantic expectancies that were carried forward to the second sentence. In this case, by the time readers encountered the word ‘explosion’ they were expecting semantic content of this sort, reducing the amount of lexico-semantic processing needed to fit the new word into their situation model. On the other hand, our framework of situation model updating is consistent with that of readers engaging in retrospective integration upon reading the second sentence. For example, upon encountering ‘explosion’ a resonance process (O’Brien, Rizzella, Albrecht, & Halleran, 1998) takes place that reactivates the features of the situation model consistent with the current lexico-semantic information, allowing for an easy fit for the new information[1].

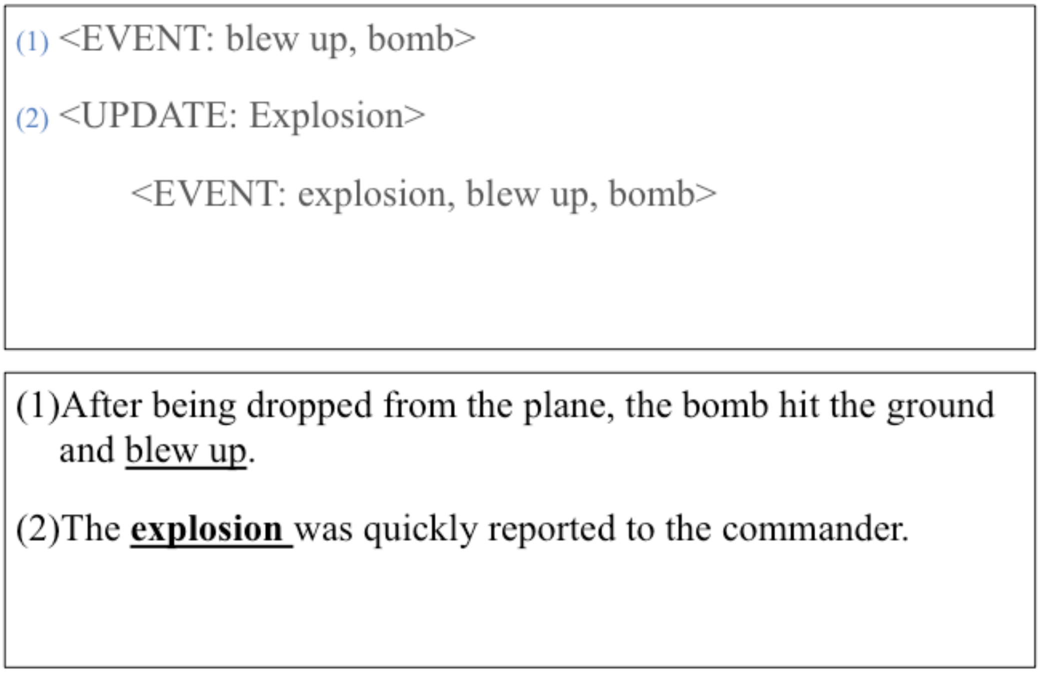

Figure 1. Simple schematic of situation model updating.

[1] It should be noted that this does not imply that resonance processes are limited to the texts with co-referential terms or events linking sentences or clauses together. We assume that resonance is a general mechanism active in comprehension, and that if this view is correct, the better semantic or situation 'match' in the paraphrase condition leads to more rapid retrieval of prior information, or to a reduced amount of information needing to be retrieved for co-referential binding, relative to the baseline condition.

1.3 LEXICAL ASSOCIATION

One way to explore the potential contributions of prospective and retrospective processes in generating the paraphrase effect found by Yang et al. (2007) is through the manipulation of lexical-level factors known to influence predictability in sparse contexts (e.g., in word pairs), and to examine the electrophysiological responses elicited by manipulating these factors in word pairs in isolation and embedded in richer contexts. As Yang and colleagues were not focused on manipulating specific lexical-level factors that may have facilitated the lexico-semantic processing of critical words in the paraphrase condition, it is not possible to tell from their study whether the results emanate mainly from prospective or retrospective processes. Thus, in the current study we manipulated a lexical-level factor known to influence to the on-line processing of words: *lexical association.*

Words that are preceded by semantically- or associatively-related words are processed more quickly and accurately than words preceded by unrelated words (Meyer & Schvaneveldt, 1971). This well-known priming phenomenon has been attributed to automatic spreading activation (ASA) at short stimulus-onset asynchronies (SOAs) and more controlled or strategic semantic expectancy at longer SOAs (Neely & Keefe, 1989). Priming effects have also been found in ERP studies, where words preceded by related words elicit reduced N400s relative to words preceded by unrelated words (Holcomb, 1988). In addition, associative priming effects have been found to be "dose dependent," as Coney (2002) found a linear decrease in primed lexical decision reaction times with increases in associative strength between prime-target pairs. The rich body of experimental evidence for these effects is supportive of the ability of individuals to predict the semantic features of the second word in associated pairs, if not the exact word itself.

In a recent study (Stafura & Perfetti, in submission), the authors manipulated the strength of forward (antecedent to critical word) associative strength across two-sentence texts adapted from the Yang et al. (2007) materials. In that study, the critical words were either strong or weak associates of the conceptually-related antecedent words in the first sentence. Electrophysiological responses measured from the onset of the critical words in the second sentence of the texts revealed an ease of word-to-text lexico-semantic fit in both strongly and weakly associated conditions, relative to baseline conditions, as indexed by reduced N400 amplitudes. Importantly, no differences in ERP responses were seen between words preceded by texts containing strong or weak associates. The authors interpreted this as indicating that, after accounting for message-level effects, there was not additional ease of processing elicited by the lexical level factor of forward association strength. Thus, in terms of the prospective account of word-to-text integration, word-level associations in the forward direction were not important in terms of ease of lexico-semantic fit. This is also consistent with research finding null or minimal effects of lexical association between words during the on-line processing of coherent texts (Coulson, Federmeier, Van Petten, & Kutas, 2005; Morris, 1994; Traxler, Foss, Seely, Kaup, & Morris, 2000; Van Petten, Coulson, Weckerly, Federmeier, Folstein, & Kutas, 1999; for effects of lexical association during online processing of coherent texts see Camblin, Gordon, & Swaab, 2007; Carroll & Slowiaczek, 1986; Hoeks, Stowe, & Doedens, 2004; Van Petten, 1993)

However, lexical association in either direction between a pair of words results in priming. Koriat (1981) was the first to document a priming effect for pairs of words that were only associated in the *target* to *prime* direction. For example, in norming tasks, a word such as 'stork' leads individuals to generate the associate 'baby' a substantial proportion of the time, but 'baby' rarely (or never) leads individuals to generate 'stork'. Nevertheless, Koriat found that, in a primed lexical decision task, pairs of words associated in either the forward (*prime* to *target*) or backward (*target* to *prime*) direction resulted in equivalent reductions in response times relative to unrelated pairs. This backward priming effect has been replicated, and such word pairs have also been shown to result in similar N400 reductions as forward associated pairs (Chwilla, Hagoort, & Brown, 1998; Peterson and Simpson, 1989). The mechanism functioning during backward priming has been suggested to be a retrospective semantic matching process (Chwilla et al., 1998; Neely & Keefe, 1989), through which processing of the target leads to a semantic match with the memory of the prime, with this co-activation leading to enhanced likelihood of classifying the target as a word (in lexical decision tasks), and a better lexico-semantic match between the words leading to a reduction in the N400 component. Thus, studies in which word pairs have been used to explore associative relations suggest a potential way to explore the relative contributions of prospective and retrospective processing by manipulating the direction of associative strength. Therefore, in this study, we manipulated the direction of association strength between pairs of words embedded in two-sentence texts, as well as word pairs used in a relatedness judgment task.

In most ERP research, including that reviewed above, it has been assumed that the N400 unitarily reflects lexico-semantic fit, whether reached through prospective or retrospective processes. Specifically, it is assumed that the N400 component does not differ in the processes it reflects by topography, at least within the broad centro-parietal scalp regions wherein it is seen when studying written words (Kutas & Federmeier, 2011). However, a growing body of ERP research using dense-array EEG nets (usually defined as consisting of 64 or more electrodes) suggests that electrophysiological responses measured at different scalp sites during the typical N400 time window (300-500 ms post stimuli) may index different processes. Dien and colleagues have suggested that several central and parietal sites can be used to explore different aspects of lexico-semantic processing in sentences (Dien, Michelson, & Franklin, 2010; Dien & O'Hare, 2008; Franklin, Dien, Neely, Huber, & Waterson, 2007). Dien et al. (2010) suggested that responses over central electrodes are associated with sequential expectancy, and are linked to earlier-occurring ERP components reflecting high-level attention processes (P2; Luck & Hillyard, 1994) and stimulus classification (P300; Johnson & Donchin, 1980; Kutas, McCarthy, & Donchin, 1977). Dien et al. also suggested that responses over parietal sites are associated with the retrospective retrieval of lexico-semantic information and its integration into its context. These 'sub-components', for lack of a better term, are suggested to combine to produce the classical N400 component, but that experimental manipulations can modulate the contribution of one or the other.

For example, Franklin and colleagues (Franklin et al., 2007) employed a primed lexical decision task to examine the possible dissociation of processes functioning during exposure to word pairs differing in association direction (e.g., prime-to-target *vs* target-to-prime). Though similar behavioral evidence of priming was found across the association conditions (i.e., shorter reaction times for associated pairs relative to unrelated pairs), the electrophysiological evidence indicated differential processing across conditions. Relative to unrelated pairs, forward and symmetrically associated prime-target pairs elicited reduced negativity over central electrodes peaking around 400 ms after target onset. However, backward associated pairs elicited a reduced negativity, relative to unrelated pairs, over right parietal electrodes. These findings are supportive of the existence of multiple priming mechanisms, with the right parietal effect suggested to index an integrative semantic matching process, and the more central effect suggested to index an expectancy process.

Finally, Dien and O'Hare (2008) used ERP and fMRI to examine sentential semantic priming effects that they attributed to automatic spreading activation. In the ERP study, the effect of semantic priming was found over left parietal electrodes. While the authors suggested that this region was responsible for automatic spreading activation, which would not be likely to occur in the current study (i.e., because of long SOAs between items), their results further support the potential for fractionating lexico-semantic effects over a relatively discreet region of scalp. That a multiplicity of processes contribute to the N400 is not surprising, given what we know about the complexity of the semantic system. And, even though localization using EEG is largely indeterminate, different patterns of activation over different clusters of electrodes is suggestive of differing neural generators, or at the very least differing amounts of neural activity.

2.0 CURRENT STUDY

Given the above background in word-to-text integration and lexical associative processing, this study examined the effects of asymmetric lexical association strength on the on-line processing of words in two tasks: text comprehension and relatedness judgments. In the text comprehension task, participants read two-sentence texts, with ERP measurements taken from a critical word in the second sentence. The experimental manipulation was the direction of strong association between the critical word and a conceptually-related paraphrase word in the preceding sentence. Participants read some texts in which the direction of strong association was from the antecedent to the critical word (*forward association* texts), and other texts in which the direction of strong association was from the critical word to the antecedent (*backward association* texts). Electrophysiological responses elicited by the critical words in the experimental texts were contrasted with those elicited during the reading of critical words in coherent control texts, wherein the words had no conceptually-related antecedent in the first sentence.

In addition to the text comprehension task, participants completed a word relatedness judgment task. The stimuli for this task were also asymmetrically associated words, matched to those collected for the comprehension task on association strength, frequency, and length. Participants made meaning relatedness decisions to pairs that were strongly forward associated, strongly backward associated, or unrelated. ERP measurements were taken from the second word of each pair. The performance on these two tasks by the same group of participants allowed us to examine lexico-semantic processing within sparse and rich contexts. In this way we attempted to fractionate ERP responses previously associated with prospective and retrospective integration.

2.1 HYPOTHESES

In terms of behavioral performance on the relatedness judgment task, consistent with previous studies (Chwilla et al., 1998; Franklin et al., 2007; Koriat, 1981), we expected to find priming effects for both types of associated pairs relative to unrelated pairs, with no differences between association conditions. In the ERP analysis of the relatedness judgments, several predictions were plausible. If the hypotheses of Franklin et al. (2007) are accurate, we expected to find a greater N400 effect (i.e., reduced negative deflection for associated pairs relative to unrelated pairs) for the forward associated condition over central electrodes, relative to the backward associated condition. In addition, we expected to find a greater N400 effect for the backward associated condition over right parietal electrodes, relative to the forward associated condition. Finally, if the effects found (albeit at an earlier point in time than the N400) by Dien and O'Hare (2008) result from more strategic expectancy processes, and not just automatic spreading activation, we expected to find a greater N400 effect for the forward associated condition over left parietal electrodes, relative to the backward associated condition.

The above predictions regarding the ERPs elicited during the relatedness judgment task assume the possibility of detecting different effects depending on scalp location, a view that differs from others on the N400 component (Kutas & Federmeier, 2011). Thus, more straightforward predictions can be made assuming similar effects across the different scalp locations. For example, if prospective processes dominate in N400 modulation, we expected a greater N400 effect for the forward association condition relative to the backward associated condition. One the other hand, if retrospective processes dominate, we expected no differences in N400 effects between the association conditions.

In the text comprehension task, a number of predictions regarding N400 responses can be made. First, if the richer context has little effect on lexico-semantic processing of the critical words, and processing differences can indeed be found across scalp sites, the same topographical differences suggested by Dien and colleages (Dien & O'Hare, 2008; Franklin et al., 2007) were expected in this task, as in the relatedness judgment task. However, if message-level context plays an important role, as is implied by previous research (Stafura & Perfetti, in prep; Yang et al., 2007), and we assume similar responses across scalp locations, contrasting predictions regarding prospective and retrospective processes can be offered. In terms of prospection, we would make a slightly different prediction than that for the relatedness judgment data. That is, because of the rich context offered by the two-sentence texts, we expected to see a similar N400 effect for both association conditions. In terms of retrospection, however, it is possible that the backward associated condition would elicit a greater N400 effect than the forward associated condition, due to the word-to-word resonance process acting on the antecedent term in the former condition.

In addition to the on-line tasks, we collected a number of off-line measures of reading ability (described in the methods section) in order to examine correlations between these and on-line performance on our tasks. Though this is a secondary aim of the current study, findings of correlations between off-line reading ability and on-line performance may help constrain interpretations of the experimental results.

2.2 METHODS

2.2.1 Participants

Thirty-one participants were recruited from the University of Pittsburgh student and staff community. All participants were right-handed, native English speakers with normal or correct-to-normal vision, without any history of head injury or epilepsy, and between the ages of 18 and 35 years old. Some participants were recruited from the Pittsburgh Adult Reading Database, which includes scores on the Nelson-Denny vocabulary and comprehension test (Nelson & Denny, 1973). Other participants were recruited through advertisements placed throughout campus locations, and completed the Nelson-Denny tests after their experimental sessions (described below). The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), which includes subtests measuring word reading efficiency and non-word decoding ability, was administered to all participants. Participants were compensated at a rate of $10 per hour, and all procedures were performed with permission from the University of Pittsburgh Institutional Review Board.

2.2.2 Materials

Word pairs were chosen using the South Florida Association Norms (Nelson, McEvoy, & Schreiber, 1998). Pairs were chosen such that the association strength was asymmetrical, i.e., strong in one direction and weak or nonexistent in the other. The association strength in the dominant direction was at least .20, and in the weak direction no pair had association strength greater than .05 (Frishkoff, 2007). For the text comprehension task, 90 word pairs were collected with mean association strength in the strong (forward) direction of .354 (SD = .14), and mean association strength in the weak (backward) direction of .017 (SD = .01). For the relatedness judgment task, 120 word pairs were collected with mean association strength in the strong (forward) direction of .348 (SD = .13), and mean association strength in the weak (backward) direction of .014 (SD = .02). The pairs did not differ across tasks in either the forward or backward directions (*p*s > .1). For the text comprehension task, the constraint of choosing pairs that fit into the contexts necessarily led to differences between the words of each pair in terms of log frequency (http://subtlexus.lexique.org/; Brysbaert & New, 2009), (mean (SD) log word freq = 2.72 (.55) and 3.39 (.54), $p < .001$) and length (mean (SD) letters = 5.9 (1.5) and 4.3 (1.3), $p < .001$). As word pairs for the relatedness judgment task were chosen to match those in the sentence comprehension task, the words in each of these pairs also differed in terms of log frequency (mean (SD) log freq = 2.64 (.56) and 3.52 (.61), $p < .001$), and length (mean (SD) letters = 5.5 (1.5) and 4.87 (1.5), $p < .001$). However, as the word pairs were seen in different orders approximately equal times across participants, effects of length and frequency differences should have been attenuated. Additionally, the pairs chosen for the relatedness judgment task did not differ from those chosen for the text comprehension task on either frequency or length (*p*s > .5).

A total of 90 two-sentence experimental texts were created for the study. The first sentence of each passage contained one asymmetrically associated word (i.e., antecedent), and the other associate (i.e., the *critical word*) was always the second word of the second sentence (i.e., the first content word). The antecedent and the critical words were chosen such that, in either direction, in the context of the passages the items are coherent. In the forward associated (FA) text condition, the strong association direction was from the antecedent word to the critical word. In the backward associated (BA) text condition, the strong association strength direction was from the critical word to the antecedent word. A control (Control) text condition was created by removing the conceptually related antecedents from the first sentences, as well as slightly changing word order to maintain coherence. The control texts were not created to be anomalous, but congruent, coherent texts. The semantic content of the control and experimental texts was compared by using the document to document tool on the Colorado University Latent Semantic Analysis website (http://lsa.colorado.edu/; Landauer & Dumais, 1997), which revealed a mean (SD) pairwise similarity metric of .804 (.15) between the conditions. Each word of the collected pairs was used as the critical word in the control texts for half of the participants. In all, four versions of each passage were created (two experimental and two control), and each version was assigned to a separate list, with the lists used approximately equally across participants. No participant saw more than one version of a given text. Examples of the passages are shown in Table 2.

Table 2. Sample Passages for Each Experimental Condition

Text Condition	Sample Passage
Forward Associated	While Cathy was riding her bike in the park, dark clouds began to gather, and it started to rain. The **storm** ruined her beautiful sweater.
Backward Associated	While Cathy was riding her bike in the park, dark clouds began to gather, and it started to storm. The **rain** ruined her beautiful sweater.
Control #1	When Cathy saw there were no dark clouds in the sky, she took her bike for a ride in the park. The **rain** that was predicted never occurred.
Control #2	When Cathy saw there were no dark clouds in the sky, she took her bike for a ride in the park. The **storm** that was predicted never occurred.

Note. The critical word (rain) is underlined and in bold at the beginning of the second sentence. The antecedent words in the paraphrase conditions are underlined.

As mentioned above, 120 word pairs were collected for use in the word relatedness judgment task. Eight lists were created by dividing the word pairs into four sets of 30 pairs each, and assigning the pairs to three different pair conditions: 30 forward associated (FA) pairs with the strong association direction from prime to target, 30 backward associated (BA) with the strong association direction from target from prime, and 60 unrelated (Unrl) pairs in which each word was paired with an unrelated word taken from the other pairs. The Unrl pairs were checked by hand to assure that they were not associated according to the USF word association database (Nelson et al., 1998). The lists were used an approximately equal number of times across participants.

2.2.3 Design and Procedure

At the beginning of the experimental session, participants were fitted with an electroencephalogram (EEG) net and seated in a sound-attenuated, electrically insulated booth. Participants were seated in an adjustable chair 60 cm from the center of a 15-in (38.1 cm) CRT display. The first ERP task was counterbalanced across participants such that an approximately equal number of participants took part in the text comprehension task first and the relatedness judgment task first. The TOWRE was administered in between the ERP tasks.

During the text comprehension task, participants read two-sentence passages for comprehension. The sentences were presented one word at the time in the center of a computer screen for a duration of 300 ms with an inter-stimulus interval (ISI) of 300 ms (i.e., stimulus-onset asynchronies (SOAs) of 600 ms). The ISI after the last word of the first sentence was increased to 600 ms to account for sentence wrap-up effects (Just & Carpenter, 1980; Raynor, Sereno, Morris, Schmauder, & Clifton, 1989). The texts were preceded by a fixation cross (+) to orient the participants. A true-false comprehension question based on the meaning of the passage followed 30% of the trials on a random basis. For each list, half of the questions required a 'true' response, and half required a 'false' response, with responses registered using a response box. The comprehension questions were used to insure that participants read for comprehension, and immediate feedback was displayed on the screen ("Wrong" in red for incorrect responses and "Good Job" in blue for correct responses). The text comprehension portion of the experimental session was broken into three blocks of trials taking approximately 15 minutes each, to allow for breaks, and the stimuli were presented in random order. Three practice texts preceded the experimental trials. An example of a sentence comprehension trial is shown in Figure 2.

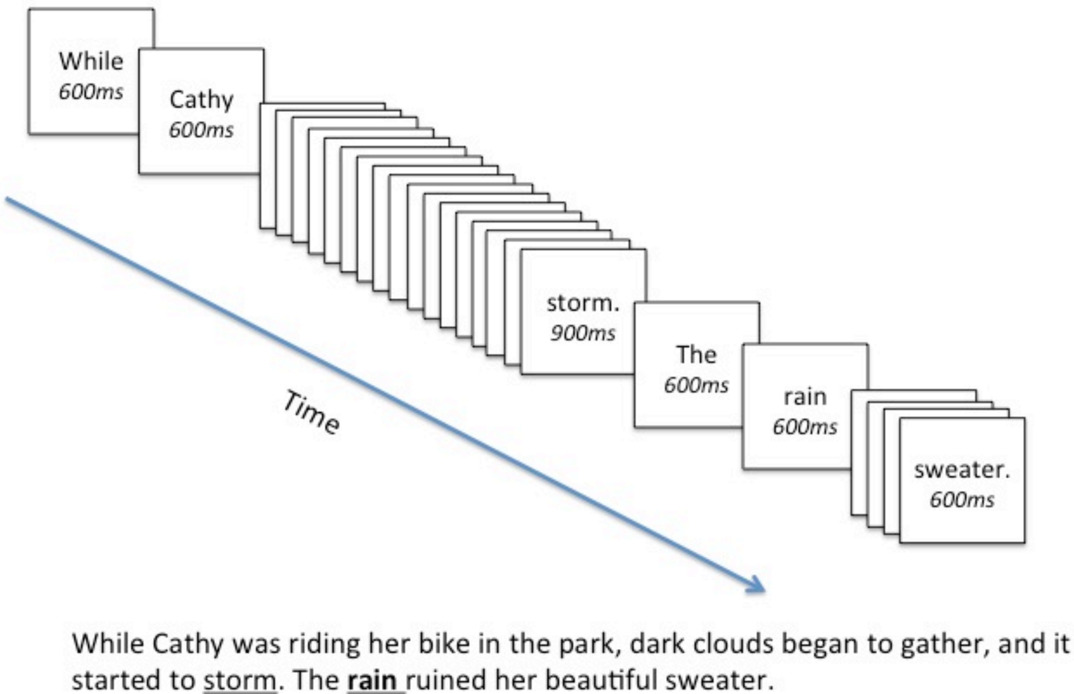

Figure 2. Example sentence comprehension trial.

During the relatedness judgment task, pairs of words were presented on the screen one at the time, and upon presentation of the second (i.e., critical) word participants were asked to respond as to whether the word was related or not related in meaning to the preceding (i.e., prime) word using a response box. The prime word was presented for 1000 ms, followed immediately by the critical word for 2000 ms. Prior to presentation of the first word a fixation cross (+) was presented in the center of the screen for 450 ms, followed by a blank screen for a random duration between 75-250 ms. During six practice trials, participants received feedback after responding (“Wrong” in red for incorrect responses and “Good Job” in blue for correct responses). During experimental trials participants did not receive feedback unless no response was registered within the 2000 ms presentation time of the critical word (“No Response” in red). The relatedness judgment portion of the experiment was broken into 3 blocks taking approximately 4 minutes each, to allow for breaks, and the stimuli were randomly presented.

In between the ERP tasks, after a short break, the TOWRE was administered. Participants remained in the booth during this task, which was audio recorded for offline scoring. The TOWRE consists of two tests of verbal fluency and decoding. In the word reading efficiency subtest participants were asked to orally read as many words as they could in 45 s from a sheet of paper consisting of 4 columns of words of increasing length and complexity (number of items = 104). In the non-word decoding subtest participants were asked to orally decode as many non-words as they could in 45 s from a sheet of paper consisting of three columns of non-words increasing in length and complexity (number of items = 63).

After completing the final ERP task, participants were lead from the EEG booth, and the electrode net was removed. For the participants who had already taken the Nelson-Denny test as part of the Pittsburgh Adult Reading Database battery, the experimental session ended at this point. For those participants recruited outside of the Database, the Nelson-Denny vocabulary and comprehension tests were administered. The Nelson-Denny vocabulary test is comprised of 100 questions that assess vocabulary knowledge, and participants were asked to complete as many questions as they could in 7.5 minutes. The Nelson-Denny comprehension test features 6 text passages followed by comprehension questions (total number of questions = 36), and participants were asked to complete as many questions as they could in 15 minutes.

2.2.4 Apparatus and ERP Recordings

During the experimental session, participants wore a 128 electrode Geodesic sensor net (Tucker, 1993) with Ag/AgCl electrodes (Electrical Geodesics, Inc., Eugene, OR). During recording, all impedances were kept below 40kΩ, an acceptable level with this system (Ferree, Luu, Russell, & Tucker, 2001). A vertex reference was used during the recording. Six eye channels were monitored to allow for rejection of ocular artifacts. The EEG signals were digitally sampled at a rate of 500 Hz, and hardware filtered during recording between 0.1 and 200 Hz. After recording, the recorded EEG was ran through a 30 Hz low-pass finite impulse response filter. For both ERP tasks, the data was then segmented from 150 ms before to 700ms after the onset of the critical words (850 ms segments). In order to keep the number of trials consistent across conditions for the relatedness judgment task, half of the 60 Unrl trials were deleted prior to artifact detection. This was accomplished by removing all even number trials for every participant so that, prior to artifact detection, each had 30 trials each of the FA, BA, and Unrl conditions. Ocular artifact detection based upon the regression technique of Gratton, Coles, and Donchin (1983), and implemented in NetStation, was used to regress out eye-blink and eye-movement behavior. Next, channels were automatically removed from the datasets if they had activity of ±200 μV, using a 80ms moving average, on more than 20% of trials. Additionally, segments were removed on the basis of three separate criteria: containing more than 12 channels marked using the previous noisy channel thresholding step, containing blinks revealed by voltage fluctuations of ±140 μV at superior and inferior eye channels (excepting for the right superior eye channel removed during ocular artifact detection), or containing horizontal eye movements (e.g., saccades) revealed by voltage fluctuations of ±55 μV at the left outer canthi electrode. Finally, individual datasets were manually checked for additional noisy channels.

The data for two participants were removed due to more than 10 trials per condition being marked bad (through the methods above) on both tasks; this data is not included in the subsequent analysis. Additionally, the data from two additional participants' text comprehension tasks were removed for the same reason. For the remaining datasets, an average of 7 electrodes was removed. The remaining trials for the different conditions in both tasks were not unbalanced. Removed channels were replaced by the data from neighboring channels using spherical spline interpolation (Ferree, 2006). The cleaned data were re-referenced to the average of the channels and corrected for the Polar Average Referencing Effect (PARE; Junghöfer, Elbert, Tucker, & Braun, 1999). The data were then averaged within participants for each condition. Following subtraction of the mean amplitude of the baseline period (150 ms pre-stimulus for both tasks), the data were exported to SPSS 19.0 for statistical analyses.

All computerized experimental tasks were programmed and carried out on E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA). E-Prime also sent event information to the EGI NetStation EEG recording system. Instructions and the computerized trials (i.e., text comprehension and relatedness judgments) were presented on a 15-in. (38.1-cm) CRT display with a 60 Hz refresh rate.

3.0 RESULTS

3.1 DESCRIPTIVE DATA

Table 3 displays participant descriptive data, along with full and partial correlations among the individual difference measures. Participants' mean Comprehension scores (25.24) were about one standard deviation above the mean of the 6328 participants in the Pittsburgh Adult Reading Database (M = 20.86; SD = 5.9). Participants' mean Vocabulary scores (64.07) were also about one standard deviation above the mean of those in the Database (M = 49.05; SD = 15.6). Participant's mean scores on the Word Reading (106.2) and Decoding (103.6) sub-tests of the TOWRE were within the average range of these standardized scores.

Table 3. Participant Descriptive Information

Participants: n=29 (Female= 17)		Full & Partial Correlations				
Variable	Mean (sd)	Variable	ND Comp [a]	ND Vocab [a]	TOWRE WR[b]	TOWRE Decoding[b]
Age	21.97 (3.7)	ND Comp [a]		.554**	.301	-.041
ND Comp [a]	25.24 (5.4)	ND Vocab [a]	.572**		.156	.322~
ND Vocab [a]	64.07 (14.1)	TOWRE WR[b]	.210	.060		-.167
TOWRE WR[b]	106.21 (10.1)	TOWRE Decoding [b]	-.231	.418*	-.173	
TOWRE Decoding[b]	103.62 (11.1)					

Note. On the left are descriptive statistics for the sample of participants in this study. On the right is a full and partial correlation matrix of the individual difference measures among the sample. Full correlations are above the diagonal. Correlations after partialling out all other individual difference measures are below the diagonal. ND Comp = Nelson-Denny Comprehension. ND Vocab = Nelson-Denny Vocabulary. TOWRE WR = TOWRE Word Reading.

[a] Scores refer to raw number of items answered correctly. [b] Standard Scores.

~ $p < .1$. * $p < .05$. $p < .01$.

The full correlations among the individual difference measures revealed a significant (r = .554, $p < .01$) correlation between Comprehension and Vocabulary, and a marginally significant correlation between Vocabulary and Decoding (r = .322, $p < .1$). In general, the decrease in the correlations between scores while partialling out the other individual difference measures revealed the shared variance among the measures. However, partialling out the TOWRE sub-tests lead to a slight increase in correlation strength between Comprehension and Vocabulary (r = .572, $p < .01$). Additionally, partialling out Comprehension and Word Reading lead to an increased correlation between Vocabulary and Decoding (r = .418, $p < .05$).

3.2 TEXT COMPREHENSION

In the text comprehension task, accuracy on the comprehension questions was above 85% across conditions. This indicated to us that participants were paying attention during the passive reading task.

To examine N400 differences across conditions in the text comprehension task, mean amplitudes from 300ms to 500 ms after the onset of the critical word were averaged from three centro-parietal clusters (Left/P3, Central/Cz, Right/P4; figure 3), each consisting of 6-7 electrodes. These electrode clusters provide broad coverage over centro-parietal areas where classic N400 effects are found (Kutas & Federmeier, 2011) and where N400 effects were seen in our previous research (Stafura & Perfetti, in prep), as well as provide coverage over scalp regions suggested to index differential kinds of processing (Dien et al., 2010; Dien & O'Hare, 2008; Franklin et al., 2007).

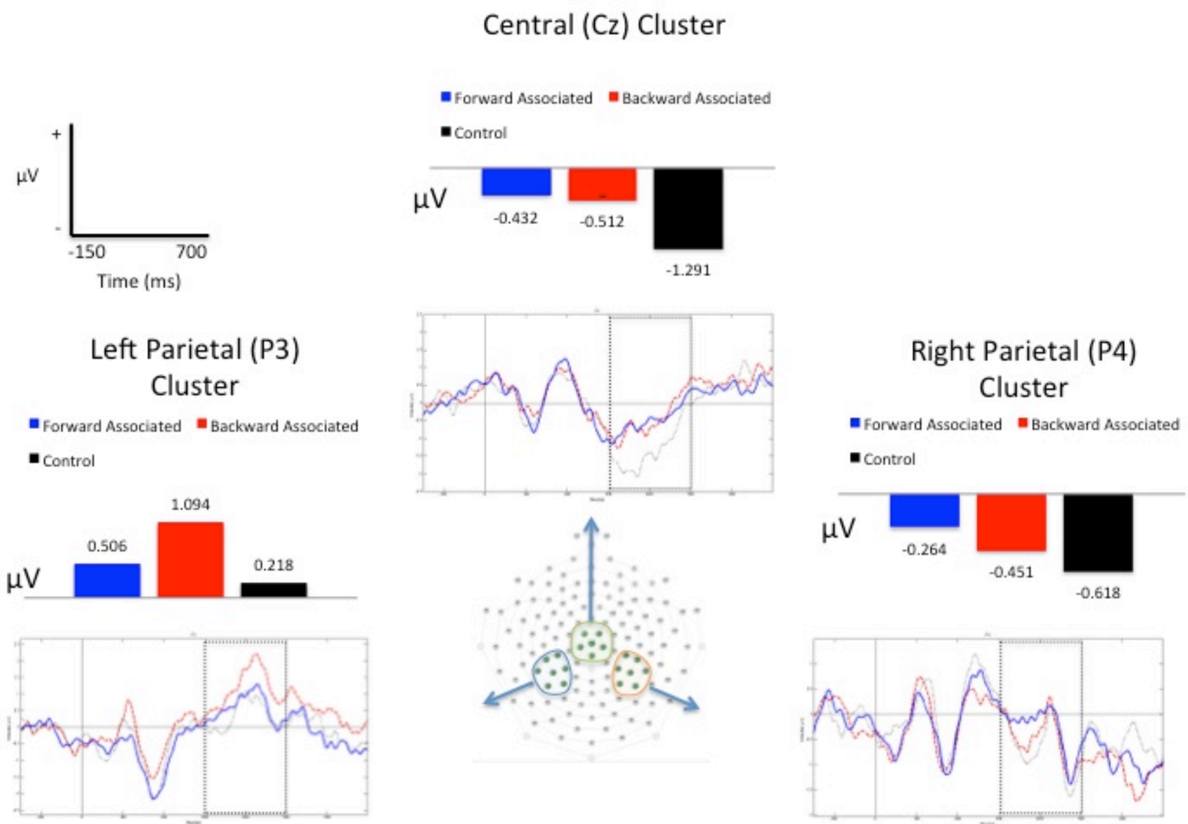

Figure 3. Waveforms and amplitude charts for the text comprehension task ERP data.

On the bottom is a schematic of the electrode net used in this study (anterior at the top), along with the three electrode clusters of interest (highlighted). On the right side of the schematic is the averaged waveform for the right (P4) parietal cluster, on the left side of the schematic is the averaged waveform for the left (P3) parietal cluster, and above the schematic is the averaged waveform for the central (Cz) cluster. The onset of the critical word is marked by the thin vertical line close to the left end of each waveform, and the 300-500ms time window of interest is indicated by the thicker black box further to the right end of the

waveforms. The averaged amplitude data (in μV) across the 300-500ms time-window for each condition, and for each cluster, is shown above each respective waveform.

Three repeated-measures analyses of variance (ANOVAs) were carried out on the amplitude data across the three Conditions (FA, BA, Control). These analyses revealed reliable Condition effects over the Cz cluster, $F(2,52) = 4.976$, $p = .011$, $\eta_p^2 = .161$, and P3 cluster, $F(2,52) = 9.224$, $p < .001$, $\eta_p^2 = .262$, but not the P4 Cluster, $F(2,52) < 1$ (see Table 4 for full statistics). For the Cz cluster, a priori comparisons revealed differences between the Association Conditions and the Control condition. Paired comparisons revealed differences between the FA and Control conditions, $t(26) = 2.494$, $p = .019$, and between the BA and Control conditions, $t(26) = 2.340$, $p = .027$, but not between the FA and BA conditions, $t(26) < 1$. These effects reflect a reduced negative deflection between 300-500ms for the FA and BA conditions relative to the Control condition, as illustrated by the waveform in Figure 3.

For the P3 cluster, a priori comparisons revealed differences between the BA Condition and the other Conditions. Paired comparisons revealed differences between the BA and Control conditions, $t(26) = 4.256$, $p < .001$, and between the BA and FA conditions, $t(26) = -2.806$, $p = .009$, but not between the FA and Control conditions, $t(26) = 1.385$, $p = .178$. These effects reflect an increased positivity between 300-500ms for the BA condition relative to the other conditions, as illustrated by the waveform in Figure 3.

Table 4. Analysis of Variance (ANOVA) on Text Comprehension Task Mean Amplitude from 300-500ms at Cz, P3, and P4 Electrode Cluster Sites

		Cz Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 52	4.976	7.83	.011*	.161
A Priori Contrasts[a]	*t(26)*	*p*			
FA *vs* BA	.387	.702			
FA *vs* Control	2.494	.019*			
BA *vs* Control	2.340	.027*			
		P3 Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 52	9.224	5.39	<.001***	.262
A Priori Contrasts[a]	*t(26)*	*p*			
FA *vs* BA	-2.806	.009**			
FA *vs* Control	1.385	.178			
BA *vs* Control	4.256	<.001***			
		P4 Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 52	.594	.910	.545	.022

Note. The Greenhouse-Geisser correction was applied when the degrees of freedom were greater than 1. The corrected *p* values and *MSE* are reported. FA = Forward Associated. BA = Backward Associated.
[a] Uncorrected.
$^{\sim} p < .1$. * $p < .05$. ** $p < .01$. *** $p < .001$

3.3 RELATEDNESS JUDGMENTS

3.3.1 Behavioral

Table 5 shows the complete behavioral data (error rates and reaction times) for the relatedness judgment task. Participants were equally likely to judge FA pairs as semantically related (96%) as they were BA pairs (95%), $t(28) = 1.00$, $p > .3$. For decision times, we compared FA and BA trials on which a 'Related' response was recorded, and Unrl trials on which an 'Unrelated' trial was recorded. Reaction times on all conditions differed reliably, all *p*s < .001. Participants responded most quickly in the FA condition (588 ms), followed by the BA condition (627 ms), with responses being the slowest in the Unrl condition (667 ms).

Table 5. Relatedness Judgment Mean (Standard Deviation) Accuracy and Reaction Times

Condition	Accuracy (proportion correct)	RT (ms)
Forward Associated	.96 (.04)	588 (82)
Backward Associated	.95 (.05)	627 (80)
Unrelated	.89 (.08)	667 (94)

Note. Accuracy refers to the proportion of 'Related' responses to Forward Associated and Backward Associated word pairs, and 'Unrelated' responses to Unrelated pairs.

3.3.2 ERP

As with the text comprehension trials, we examined the mean amplitudes from 300 to 500 ms from the onset of the critical (second) word in the relatedness judgment trials. The same three centro-parietal clusters (P3, CZ, P4; figure 4) used for the text comprehension analyses were used for these analyses.

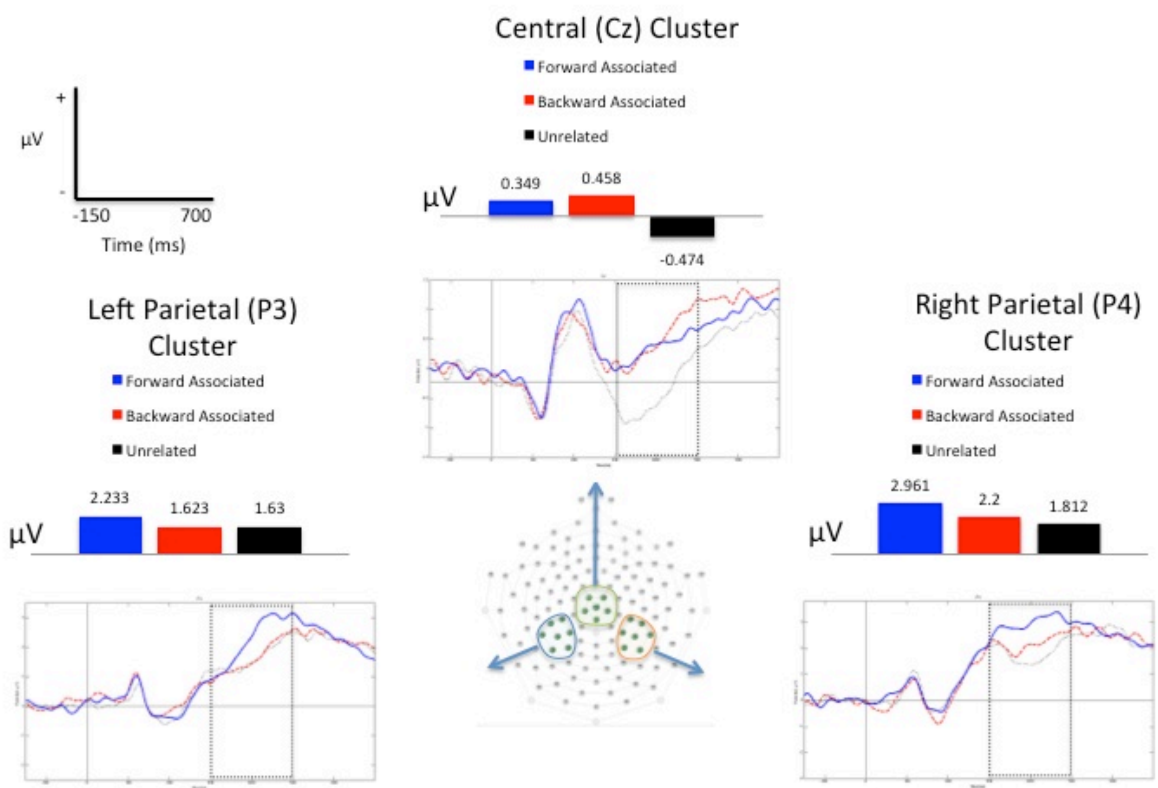

Figure 4. Waveforms and amplitude charts for the relatedness judgment task ERP data.

On the bottom is a schematic of the electrode net used in this study (anterior at the top), along with the three electrode clusters of interest (highlighted). On the right side of the schematic is the averaged waveform for the right (P4) parietal cluster, on the left side of the schematic is the averaged waveform for the

left (P3) parietal cluster, and above the schematic is the averaged waveform for the central (Cz) cluster. The onset of the critical word is marked by the thin vertical line close to the left end of each waveform, and the 300-500ms time window of interest is indicated by the thicker black box further to the right end of the waveforms. The averaged amplitude data (in µV) across the 300-500ms time-window for each condition, and for each cluster, is shown above each respective waveform.

Three repeated measures analyses of variance (ANOVAs) were carried out on the amplitude data across the 3 Conditions (FA, BA, Unrl). These analyses revealed reliable Condition effects over the Cz cluster, $F(2,56) = 7.593$, $p < .001$, $\eta_p^2 = .213$, and P4 cluster, $F(2,56) = 7.318$, $p = .002$, $\eta_p^2 = .207$, electrode clusters, but only a marginal effect over the P3 electrode cluster, $F(2,56) = 2.929$, $p = .075$, $\eta_p^2 = .095$ (see Table 6 for full statistics). For the Cz cluster, a priori comparisons revealed differences between the Association Conditions and the Unrl condition. Paired comparisons revealed differences between the FA and Unrl conditions, $t(28) = 3.179$, $p = .004$, and between the BA and Unrl conditions, $t(28) = 3.133$, $p = .004$, but not between the FA and BA conditions, $t(28) < 1$. These effects reflect a reduced negative deflection between 300-500ms for the FA and BA conditions relative to the Unrl condition, as illustrated in the waveform in Figure 4.

For the P4 cluster, a priori comparisons revealed differences between the FA Condition and the other Conditions. Paired comparisons revealed differences between the FA and Unrl conditions, $t(28) = 3.431$, $p = .002$, and between the FA and BA conditions, $t(28) = 3.134$, $p = .004$, but not between the BA and Unrl conditions, $t(28) = 1.175$, $p = .250$. These effects reflect a reduced negative deflection between 300-500ms for the FA condition relative to the other conditions (Figure 4). Though the Condition effect over the P3 cluster was only marginal ($p = .075$), the direction of effects was consistent with those seen over the P4 cluster.

Table 6. Analysis of Variance (ANOVA) on Relatedness Judgment Task Mean Amplitude from 300-500ms at Cz, P4, P3 Electrode Cluster Sites

		Cz Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 56	7.593	8.339	.001**	.213
A Priori Contrasts[a]	*t(28)*	*p*			
FA *vs* BA	-.491	.627			
FA *vs* Unrl	3.179	.004**			
BA *vs* Unrl	3.133	.004**			
		P4 Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 56	7.318	11.253	.002**	.207
A Priori Contrasts[a]	*t(28)*	*p*			
FA *vs* BA	3.134	.004**			
FA *vs* Unrl	3.431	.002**			
BA *vs* Unrl	1.175	.250			
		P3 Cluster			
	df	*F*	*MSE*	*p*	η_p^2
Condition	2, 56	2.929	4.505	.075~	.095

Note. The Greenhouse-Geisser correction was applied when the degrees of freedom were greater than 1. The corrected *p* values and *MSE* are reported. FA = Forward Associated. BA = Backward Associated. Unrl = Unrelated.

[a] Uncorrected.

~ $p < .1$. * $p < .05$. ** $p < .01$. *** $p < .001$.

3.4 INDIVIDUAL DIFFERENCE CORRELATIONS

3.4.1 Behavioral

As can be seen in Table 7, several reaction time measures from the relatedness judgment task were correlated with the individual difference measures. Comprehension scores were negatively correlated (*p*s < .05) with the time to make judgments that two backward associated words were related semantically, and the time to make judgments that two unrelated words were unrelated semantically. The correlation between Comprehension scores and Unrl decision speed led to negative correlations between reaction time differences between Unrl decisions and related decisions for both FA and BA words pairs (*p*s < .05). This negative association was also found for Vocabulary scores (*p*s < .05). Finally, the difference between relatedness judgments for BA word pairs and FA word pairs was negatively correlated with Word Reading scores ($p < .05$).

Table 7. Relatedness Judgment Reaction Time Data and Individual Difference Correlations

	Nelson-Denny Comprehension	Nelson-Denny Vocabulary	TOWRE Word Reading
BA	-.386*		
Unrl	-.446*		
Unrl - FA	-.409*	-.458*	
Unrl - BA	-.461*	-.501**	
BA - FA			-.392*

Note. Only correlations of $p < .05$ are shown. BA = Backward Associated. FA = Forward Associated. Unrl = Unrelated. Unrl-FA = Unrelated minus Forward Associated. Unrl-BA = Unrelated minus Backward Associated. BA-FA = Backward Associated minus Forward Associated.

* $p < .05$. ** $p < .01$.

3.4.2 ERP

Table 8 shows correlations between ERP measures and reading skill measures. For the relatedness judgment data, correlations were found with ERP measures over the central (Cz) and left (P3) parietal clusters. Over the central cluster, Comprehension scores were positively correlated with N400 amplitude for FA word pairs ($p < .05$), which likely resulted in the marginal correlation between Comprehension scores and the amplitude difference between FA and Unrl trials ($p < .1$). Additionally, a marginal negative correlation was found between Vocabulary scores and amplitude in response to BA pairs over the left (P3) Parietal cluster ($p < .1$), which likely resulted in the marginal positive correlation between Vocabulary scores and the amplitude difference between FA and BA trials ($p < .1$).

Table 8. Event-Related Potential (ERP) Mean Amplitude from 300-500ms and Individual Difference Correlations

	Nelson-Denny Comprehension	Nelson-Denny Vocabulary
RJ Cz FA	.390*	
RJ Cz FA - Unrl	.354~	
RJ P3 BA		-.356~
RJ P3 FA - BA		.322~
Text Cz Control	.513**	
Text Cz FA - Control	-.358~	

Text Cz	
BA - Control	-.442*

Note. Only correlations of $p < .1$ are shown. RJ = Relatedness Judgment. Text = Text Comprehension. Cz = Central electrode cluster. P3 = Left parietal cluster. FA = Forward Associated. BA = Backward Associated. Unrl = Unrelated. FA-Unrl = Forward Associated minus Unrelated. FA-BA = Forward Associated – Backward Associated. FA- Control = Forward Associated minus Control. BA-Control = Backward Associated minus Control.
~ $p < .1$. * $p < .05$. ** $p < .01$.

For the text comprehension task, reliable correlations were found with ERP measures over the central (Cz) cluster. Comprehension scores were positively correlated with amplitude for the Control condition ($p < .01$), which likely resulted in the significant negative correlation between Comprehension scores and the amplitude difference between BA and Control trials ($p < .05$), and the marginal negative correlation between Comprehension scores and the amplitude difference between FA and Control trials ($p < .1$).

4.0 DISCUSSION

This study explored whether the manipulation of a lexical-level variable, the direction of word association, would allow for observable electrophysiological dissociations between prospective and retrospective integration processes during the course of processing word pairs in isolation and in texts. The N400 component, a negative deflection of the ERP waveform occurring between 300 and 500 ms after exposure to any potentially meaningful stimulus, was used to index lexico-semantic processing in a text comprehension task, and a word relatedness judgment task. We explored the effect of our lexical association manipulation at scalp sites that have been suggested to index separate lexico-semantic processes that contribute to the N400 component (Dien et al., 2010; Dien & O'Hare, 2008; Franklin et al., 2007).

In both the text comprehension task and the relatedness judgment task, N400 responses over central electrodes indicated that word pairs associated in either the forward and backward direction lead to facilitated lexico-semantic processing. The effect for the backward associated pairs suggests that responses at this scalp site do not only reflect prospective processing. Additionally, due to the differing degrees of expectancy available for word pairs associated in forward and backward directions, this result provides evidence that the binding operations in text comprehension processes described at the message level in previous research (Stafura & Perfetti, in prep; Yang et al., 2007) can be supported by retrospective processes. While prospective expectancy processes may also play a role, this is contrary to strong lexical access views of the N400 component (Lau, Almeida, Hines, & Poeppel, 2009; Lau, Phillips, & Poeppel, 2008), as association strength is defined by the differing amount of prediction of one word given the presentation of another. Since people are far less likely to predict the second word from the first in the case of the backward association condition, compare to the forward association condition, it is difficult to reconcile this with a strong view of lexical prediction. These findings are also inconsistent with the suggestion by Dien and colleagues (2010) that responses over this area index general sequential expectations of upcoming items.

Though the N400 responses were similar across the two tasks at a central scalp site, they differed at the other sites from which responses were measured. In the relatedness judgment task, relative to unrelated word pairs, a reduction in N400 amplitude was found only for forward associated pairs over right parietal electrodes (with the same pattern visible over left parietal sites). Our results differ from those of Franklin et al.'s (2007), where an N400 reduction for backward associated pairs, relative to forward associated pairs, was found over this region. This may be due to task differences, as Franklin et al. had their participants perform a lexical decision task, while our study utilized a relatedness judgment task. Judging the semantic relatedness of two terms requires accessing and contrasting both words' meanings, while making a lexical decision judgment may be possible with a shallower level of processing of the target word. The Franklin et al. findings may reflect retrospective semantic matching engaged by backward associated pairs, which would result in a greater degree of semantic activation for the target word than that seen for forward associated pairs. So while it is potentially the case that responses over this right parietal site index expectancy to a greater degree than retrospective processing, it may also be the case that they index greater lexico-semantic activation overall.

In the text comprehension task the results over parietal sites differed from those in the relatedness judgment task in somewhat unexpected ways. Over right parietal electrodes no effect of condition was found. Taking the argument from the previous paragraph, this might have resulted from a similar amount of lexico-semantic activation across the text conditions. Potentially, the rich contexts, coupled with the overall good comprehension skills among our participants, made it relatively easy to process the control texts, which were not made to be incongruent in any way. This perspective would seem to be inconsistent with the idea of expectancy processes at this scalp location, as the experimental texts should have, in theory, constrained the semantic features expected in the second sentences, given the co-referential terms in the first sentences, to a greater degree than the control texts. However, as the waveform in Figure 3 illustrates, there is a visible difference between the forward associated condition and the other conditions, in the same direction as that seen in the relatedness judgment data. This attenuated difference would be consistent with a message level expectancy effect driven by the contexts in the backward association and control conditions, though no definite evidence is provided by this study.

In addition, the waveforms seen over left parietal electrodes in the text comprehension task had a different structure than those seen in the relatedness judgment data. First, condition effects were found to the effect that there was a greater positivity for the backward associated condition relative to the other conditions. Stating this as a greater positivity, rather than a reduced negativity, seems appropriate, given the clear lack of a negative deflection in the 300-

500 ms time window. One conclusion is that the activation over left parietal sites reflects the positive end of a dipole summating underneath the central and right parietal regions. This cannot be confirmed (or disconfirmed) from the data at hand, though the different order of condition effects seen across the sites hints that this may not be the case, especially when contrasted with the relatedness judgment data.

Another, more speculative interpretation is that the positivity seen over the left parietal sites reflects an overlaying of positive ERP components with the N400 component. One potential positive component is the P300, a positive going deflection of the waveform originally associated with categorization of stimuli in oddball paradigms (Johnson & Donchin, 1980). The psychological mechanism behind the P300 is suggested to be that of context updating in working memory (Donchin & Coles, 1988), which, if applicable to our data, may result from the mapping of novel semantic items into the situation model constructed from the previous context. While the P300 usually has an earlier onset than that seen here, it has been suggested that its latency may be delayed depending on the degree of difficulty of categorization (Kutas et al., 1977). Since lexico-semantic activation is thought to occur at the later time point then lower-level perceptual activation (e.g., as seen in simple oddball paradigms), this might be the case.

A second positive component that might be occurring during our time window of interest is the semantic P600 (Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). This positivity, which is distinct from the syntactic P600 (Friederici, Hahne, & Mecklinger, 1996), may be related to the left parietal Old/New effect seen in memory studies (Rubin, Van Petten, Glisky, & Newberg, 1999). The interpretation for our data would be something like the following. Upon encountering the first content word of the new sentence (i.e., the critical word), episodic resonance processes

(O'Brien et al., 1998) are engaged leading to a co-activation of the newly encountered semantic information and the co-referential information in the readers' situation models. This happens for all text conditions, as indicated by the positive deflection in all conditions. For the forward associated and control texts, this resonance process leads to the reactivation of the propositional structure of the first sentence, consisting of chunked information in memory. For the backward associated condition, however, the backward association between words leads to reactivation of the propositional structure of the first sentence, but also of the antecedent word itself. This may lead to a breaking apart of the propositional structure in a similar way as the re-parsing processes suggested to underlie syntactic (Friederici et al., 1996) and semantic (Kuperberg et al., 2003) P600 effects. The lack of such positivity in the relatedness judgment data is consistent with the lack of P600 effects seen among words out of sentence contexts (Munte, Heinze, & Mangun, 1993). This interpretation is speculative, as noted above. Additional research manipulating the amount of information likely retrieved during sentence processing will be valuable in verifying or rejecting this hypothesis.

Finally, there were some suggestive but theoretically interesting correlations between off-line reading measures and behavioral and ERP responses during our on-line tasks. On the relatedness judgment task, comprehension skill was associated with reduced latencies on judgments for backward associated and unrelated word pairs. This may reflect better semantic processing abilities of more skilled comprehenders (Nation & Snowling, 1999), which allows them to rapidly categorize pairs of words based on common meanings, or lack thereof. The association between reaction times and vocabulary, which is moderately correlated with comprehension, is also consistent with this interpretation. In the relatedness judgment ERP data, comprehension skill was positively associated with N400 amplitude in response to forward associated pairs over central electrodes. In the text comprehension ERP data, over this same electrode site, comprehension skill was positively associated with N400 amplitude elicited by the control texts. Both of these results are consistent with greater semantic processing abilities among the good comprehenders. For the relatedness judgments, this may have resulted in a more rapid lexico-semantic match between the forward associated words. For the text comprehension, the results for the control condition may reflect the ease with which new lexico-semantic information can be integrated into skilled comprehenders' situation models, due to their higher quality lexical representations (Perfetti & Hart, 2002; Perfetti, 2007), which allow for rapid retrieval of context-relevant semantic features, or to domain general processes such as increased working memory capacity (Just & Carpenter, 1992), which allow for the manipulation of a greater amount of information at a time.

The results from this study support the idea that retrospective processes, and not just prospective expectancy processes, modulate N400 responses (Brown & Hagoort, 1993; Yang et al., 2007). This can be seen most clearly over central scalp sites, where across tasks words associated in both directions led to reduced negative ERP deflections relative to contexts lacking such lexical association. The different findings at parietal sites, at least for the relatedness judgments, are consistent with recent suggestions that topographical differences during the typical N400 time-window can differential processing mechanisms (Dien et al., 2010; Franklin et al., 2007), though our results seem inconsistent with the previous interpretations of ERP findings. This dissociation of processes will need to be tested using different materials and methods, but could provide a valuable new avenue of research into an electrophysiological response that has yielded three decades of important psychological and psycholinguistic findings (Kutas & Federmeier, 2011). Finally, the positivity seen over left parietal sites during passive text reading will be important to examine further. While this effect may be an artifact of measuring cortical activity at the scalp (e.g., dipolar activity), it may provide a way to examine differing amount of memory retrieval and updating, driven by resonance processes engaged during reading (O'Brien et al., 1998). Overall, this study provides evidence for the functioning of retrospective integration processes in word-to-text integration, likely in addition to prospective processes. Hopefully, this set of findings, coupled with previous work (Stafura & Perfetti, in prep; Yang et al., 2007) will spur ERP research on higher-level discourse processes, and their connection to sentence level processes, and word level processes, as well.

BIBLIOGRAPHY

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

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

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5(1)*, 34-44.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavioral Research Methods, 41(4)*, 977-990.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56(1)*, 103-128.

Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14,* 509-522.

Chwilla, D. J., Hagoort, P., Brown, C. M. (1998). The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *The Quarterly Journal of Experimental Psychology Section A, 51(3)*, 531-560.

Coney, J. (2002). The effect of associative strength on priming in the cerebral hemispheres. *Brain and Cognition, 50*, 234-241.

Coulson, S., Federmeier, K.D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(1)*, 129-147.

Dien, J., Michelson, C. A., & Franklin, M. S. (2010). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: Cognitive and neuroanatomical implications. *Brain Research, 1355*, 126-140.

Dien, J. & O'Hara, A. J. (2008). Evidence for automatic sentence priming in the fusiform semantic area: Convergent ERP and fMRI findings. *Brain Research, 1243*, 134-145.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357-374.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory & Language, 41*, 469-495.

Ferree, T.C. (2006). Spherical splines and average referencing in scalp electroencephalography. *Brain Topography, 19(1-2)*, 43-52.

Ferree, T.C., Luu, P., Russell, G.S., & Tucker, D.M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Journal of Clinical Neurophysiology, 112*, 536-544.

Franklin, M.S., Dien, J., Neely, J.H., Huber, E., & Waterson, L.D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology, 118*, 1053-1068.

Friederici, A. D., Hahne, A., Mecklinger, A. (1996). Temporal structure of syntactic parsing; Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1219-1248.

Frishkoff, G.A. (2007). Hemispheric differences in strong versus weak semantic priming: Evidence from event-related brain potentials. *Brain and Language, 100*, 23-43.

Graesser. A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifacts. *Electroencephalography and Clinical Neurophysiology, 55*, 468-484.

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

Holcomb, P. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*, 66-85.

Johnson Jr., R., & Donchin, E. (1980). P300 and stimulus categorization: two plus one is not so different from one plus one. *Psychophysiology, 17*, 167-178.

Johnson-Laird, P.N. (1981). Comprehension as the construction of mental models. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 295(1077),* 353-374.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Junghöfer, M., Elbert, T., Tucker, D. M., & Braun, C. (1999). The polar average reference effect: A bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology, 110*, 1149-1155.

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

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition, 9*, 587-598.

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

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

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

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

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology, 69*, 218-233.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-240.

Landi, N., & Perfetti, C.A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language, 102(1)*, 30-45.

Lau, E. F., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for N400 context effects: Evidence from MEG. *Brain & Language, 111*, 161-172.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Natura Reviews Neuroscience, 9(12)*, 920-933.

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

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 92-103.

Munte, T. E., Heinze, H. J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience, 5*, 335-344.

Nation, K., & Snowling, M.J. (1999). Semantic processing and the development of word recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*, 85-101.

Neely J.H. & Keefe D.E. (1989). Semantic context effects on visual word processing: a hybrid prospective/retrospective processing theory. In: Bower GH, editor. *The psychology of learning and motivation: advances in research and theory (24)*, (pp. 207-248). New York: Academic Press.

Nelson, M.J., & Denny, E.C. (1973). *The Nelson-Denny Reading Test*. Boston: Houghton Mifflin.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* http://www.usf.edu/FreeAssociation/

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 1200-1210.

Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11(4)*, 357-383.

Perfetti, C.A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.). *Precursors of functional literacy* (pp.189-213). Amsterdam/Philadelphia: John Benjamins.

Peterson, R. R., & Simpson, G. B. (1989). Effect of backward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(6)*, 1020-1032.

Rayner, K., & Clifton, C., Jr. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology, 80*, 4-9.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*, 21-49.

Rubin, S. R., Van Petten, C., Glisky, E. L., Newberg, W. M. (1999). Memory conjunction errors in younger and older adults: Event-related potential and neuropsychological evidence. *Cognitive Neuropsychiatry, 16*, 459-488.

Stafura, J. Z., & Perfetti, C. A. (in submission). Word-to-text integration: Message-level and lexical level influences in ERPs.

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (1999). *Test of word reading efficiency.* Austin, TX: Pro-Ed.

Traxler, M.J., Foss, D.J., Seely, R.E., Kaup, B., & Morris, R.K. (2000). Priming in sentence processing: Lexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29(6)*, 581-595.

Tucker, D.M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology, 87*, 154-163.

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

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event related potentials. *Language and Cognitive Processes, 8*, 485-531.

Van Petten, C., Coulson, S., Weckerly, J., Federmeier, K. D., Folstein, J., & Kutas, M. (1999). Lexical association and higher-level semantic context: An ERP study. *Journal of Cognitive Neuroscience Supplement,* 46.

Yang, C.L., Perfetti, C.A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(1)*, 55-89.

Zwaan, R.A., & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Review, 123(2),* 162-185.