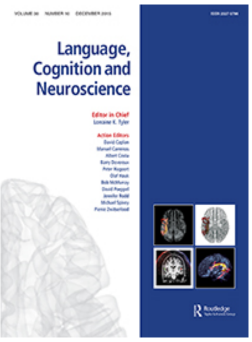

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

ERP evidence for memory and predictive mechanisms in word-to-text integration

Joseph Z. Stafura, Benjamin Rickles & Charles A. Perfetti

To cite this article: Joseph Z. Stafura, Benjamin Rickles & Charles A. Perfetti (2015) ERP evidence for memory and predictive mechanisms in word-to-text integration, Language, Cognition and Neuroscience, 30:10, 1273-1290, DOI: 10.1080/23273798.2015.1062119

To link to this article: http://dx.doi.org/10.1080/23273798.2015.1062119

Language, Cognition and Neuroscience, 2015
Vol. 30, No. 10, 1273–1290, http://dx.doi.org/10.1080/23273798.2015.1062119

ERP evidence for memory and predictive mechanisms in word-to-text integration

Joseph Z. Stafura[a,b]*, Benjamin Rickles[a] and Charles A. Perfetti[a,b]

[a]*Department of Psychology, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA;* [b]*Center for the Neural Basis of Cognition, Pittsburgh, PA, USA*

(*Received 13 November 2014; accepted 5 June 2015*)

During reading, word-to-text integration (WTI) proceeds quickly and incrementally through both prediction and memory processes. We tested predictive and memory mechanisms with event-related potentials (ERPs) recorded on critical words that were across a sentence boundary from co-referential words that differed in dominant direction of lexical association. For comparison of text comprehension, participants performed meaning judgements on a matched set of word pairs. In both tasks, reduced N400 amplitudes were elicited over central scalp electrodes by words associated in either direction relative to task-specific baseline conditions. A temporal principal component analysis of the ERP data extracted a component reflecting this central N400. Additionally, for the text comprehension task early (N200) and late (parietal N400 and P600) discriminated between forward associated and backward associated conditions. The results demonstrate that, beyond N400 indicators of prediction, ERPs reflect the role of memory processes in WTI across sentences.

Keywords: text comprehension; integration; lexical association; meaning judgements; ERPs

Readers process words, to the extent possible, as they are encountered. This view of reading as an incremental process is supported by the immediate influence of message-level factors on word-level self-paced reading (Boland, Tanenhaus, Garnsey, & Carlson, 1995; Tyler & Marslen-Wilson, 1977), eye movements (Altmann & Kamide, 1999; Rayner & Clifton, 2009), and event-related potentials (ERPs; Kutas, Van Petten, & Besson, 1988). In order to achieve such rapid processing of text, readers rely on both prospective processes that prepare for (i.e. anticipate) upcoming information and retrospective processes that link to memory for the preceding text. These forward- and backward-looking processes link word reading with text comprehension (TC), or what we refer to as word-to-text integration (WTI). The study reported here was designed to expose both forward and backward processes in WTI by measuring ERPs on critical words separated from associated antecedents by a sentence boundary.

We have previously used WTI (Perfetti & Stafura, 2015; Perfetti, Yang, & Schmalhofer, 2008; Yang, Perfetti, & Schmalhofer, 2007) to refer to the implicit meaning processes that occur when readers integrate word-level meanings into their mental models (Johnson-Laird, 1981, 1983), or situation models (Van Dijk & Kintsch, 1983), of a text. For example, Yang et al. (2007) had participants read two-sentence texts such as (1a) below.

(1a) After being dropped from the plane, the bomb hit the ground and exploded. The **explosion** was quickly reported to the commander.

(1b) After being dropped from the plane, the bomb hit the ground and blew up. The **explosion** was quickly reported by the commander.

ERPs were measured on a critical word in the second sentence, here "explosion". When readers encounter the critical word in (1a), there is both form-based and meaning-based overlap between the word and an antecedent "exploded" in the first sentence. A schematic of the potential mental model constructed during the reading of the critical word in (1a) includes both situation and event structures.

< Situation: dropped, hit ground, exploded, bomb >
< Event: Explosion >

Assuming the reader has a memory for the text and access to the situation model, the explicitly overlapping critical word is easily integrated into the situation model.

< Situation: … ,*Explosion*, exploded, bomb >

The ease of integration in response to the critical word in (1a) was reflected by a reduced N400 component elicited by critical words in this *explicit* condition relative to critical words in a *baseline* condition illustrated in (1c).

(1c) 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.

*Corresponding author. Email: jzs48@pitt.edu

The text in (1c) is sensible, not anomalous, but instead of encouraging integration with the prior text, it requires the reader to establish a new event structure (< Event: Explosion >) in their situation model (Gernsbacher, 1990, 1997), leading to an increased integration cost indexed by the N400 component. This negative-going ERP component occurs in response to any potentially meaningful stimulus, and is sensitive to the semantic match between the currently processed stimulus and its context, with smaller amplitude deflections reflecting a better match (Kutas & Hillyard, 1980; for review, see Kutas & Federmeier, 2011).

Text (1b) above, termed the *paraphrase* condition by Yang et al. (2007), differs from the explicit condition in that the critical word does not share form overlap with the antecedent, nor is it a synonym. However, due to memory of the text, the critical word "explosion" can be integrated with the situation model (which contains a "blew-up" event) through implicit WTI processes. This was reflected in a reduced N400 in the paraphrase condition relative to the baseline condition, as well as by comparable N400 responses to the explicit condition. The WTI processes engaged during reading of the critical word in the paraphrase condition depend on the meaning of the word and the meaning of the immediately preceding text. The integration processes use the context to establish a referential meaning for the word, adding it to the mental representation of the text. In the baseline texts, rather than integration, reading the critical word introduces a new event structure.

In order to test the relative influence of message-level and lexical-level factors in WTI, Stafura and Perfetti (2014) manipulated the strength of forward (antecedent to critical word) lexical associative strength across two-sentence texts. Critical words were either strong (2a) or weak (2b) associates of the referentially related antecedent words in the first sentence.

(2a) 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.

(2b) While Cathy was riding her bike in the park, dark clouds began to gather, and it started to shower. The **rain** ruined her beautiful sweater.

The ERP responses elicited by the critical words (in the above example, "rain") were recorded. For both the strongly associated and weakly associated texts, there were reduced N400 amplitudes relative to baseline texts (2c).

(2c) 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.

Importantly, no differences in ERP responses were seen between words preceded by texts containing strong compared with weak associates. Stafura and Perfetti (2014) interpreted this as indicating that, after accounting for message-level effects, the integration process was not further facilitated at the lexical level by forward word association strength. Although some studies have found effects of lexical association in coherent texts (Camblin, Gordon, & Swaab, 2007; Carroll & Slowiaczek, 1986; Hoeks, Stowe, & Doedens, 2004; Morris, 1994; Van Petten, 1993), other studies have found minimal or null effects (Coulson, Federmeier, Van Petten, & Kutas, 2005; Traxler, Foss, Seely, Kaup, & Morris, 2000). In a review, Ledoux, Camblin, Swaab, and Gordon (2006) point out that large lexical priming effects are most commonly observed in simple, or incongruous contexts, and that message (or discourse) level factors (e.g. congruence) can strongly attenuate or eliminate lexically driven N400 effects. The findings from Stafura and Perfetti (2014) are supportive of the prominence of message-level factors during comprehension of relatively rich, congruent texts.

The importance of message-level influences on WTI raises the question of mechanisms that produce these influences. Using the message-level meaning requires a memory for the text meaning and for the most recently read text segment (a clause or sentence); while effortful retrieval certainly occurs during reading, less-effortful memory processes likely play a large role. Relatively passive memory processing might function through a resonance mechanism in which the currently read word automatically activates (or increases activation of existing) links to information accessible in memory (Albrecht & O'Brien, 1993; Myers & O'Brien, 1998), in a sense acting as a retrieval cue (Ericsson & Kintsch, 1995; Ratcliff & McKoon, 1988). Integration may be facilitated when the associations of an encountered word resonate with memory of the text, which includes at least words and their meaning features, and may also include the referentially specified meaning of a situation model. If so, then *backward* association from the word being read to words (or their referents) in text memory may be functional in integration. To the authors knowledge, the studies that have examined online lexical effects in sentence and discourse processing have not controlled for, or experimentally manipulated, backward lexical association.

Lexical association in either direction between a pair of words results in priming. Koriat (1981) documented a priming effect in lexical decisions for pairs of words that were only associated in the backward *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". Koriat reported that priming in either the forward (*prime* to *target*) or backward (*target* to *prime*) direction resulted in equivalent reductions in response times relative to unrelated word pairs. Backward-associated (BA) word pairs have since been shown to elicit N400 reductions similar to those elicited by forward-associated (FA) pairs in lexical decisions (Chwilla, Hagoort, & Brown, 1998; Dien, Franklin, & May, 2006), and smaller N400 responses to unrelated words, but greater N400 responses to strongly FA pairs, in a semantic judgement task (Kandhadai & Federmeier, 2010). BA priming effects were not found in naming (Seidenberg, Waters, Sanders, & Langer, 1984), nor in cross-modal priming in which a sentence-final prime word was followed by a visually presented (BA) target word in either lexical decision or naming tasks (Peterson & Simpson, 1989).

The tripartite model proposed by Neely and colleagues (Neely, 1991; Neely & Keefe, 1989) provides a perspective on priming. In this model, forward lexical priming results from *automatic spreading activation* across lexical or semantic nodes or through controlled *strategic, expectancy processes* altering the starting landscape of activation in the lexical-semantic network. In terms of lexical decisions, backward priming likely functions through a *lexical-semantic matching process* in which the co-activation of a target and the prime facilitated the decisions that the target is a word, because only a word would have such a relationship with the prime. Thus, backward priming should not be expected when the target–prime relationship is irrelevant to the task (Forster, 1979), as it is in naming (Seidenberg et al., 1984). However, text processing provides a different situation, one in which retrospective processes (not exactly "backward priming") could be triggered by the interaction of text memory with word reading. These retrospective processes may be available to the resonance mechanism suggested to be important for comprehension (Albrecht & O'Brien, 1993).

Given the role of such memory processes as well as expectancy processes during text reading, the present study aimed to examine prospective and retrospective processing in online WTI. We did this by comparing the effects of forward association with backward association on the ERP response to a critical word. Specifically, we created two-sentence texts that contained two words, an antecedent word in sentence 1 and an associated word in the first phrase of sentence 2. The association strength was asymmetrical: Sometimes strong in the forward direction, sometime strong in the backward directions (Table 1; Supplementary table A).

Participants read some texts in which the direction of strong association (i.e. the asymmetrical association direction) 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 baseline texts, wherein the words had no co-referential antecedent in the first sentence. The ERP responses in the TC task were examined for N400 results. Additionally, The ERP data was subjected to a principal components analysis (PCA) as a data-driven approach for fractionating the ERP activation time-course (Dien & Frishkoff, 2005).

In terms of the N400 component, we expected to find an effect of association; that is, both association conditions will elicit reduced N400 amplitudes compared to the baseline condition. This is consistent with effects of prospective message-level support provided by preceding text on the integration of co-referential terms that fine-tune mental representations (Ditman, Holcomb, & Kuperberg, 2007; Otten & Van Berkum, 2008; Stafura & Perfetti, 2014; Van Berkum, Brown, & Hagoort, 1999). If the backward association between the critical word and its antecedent triggers integration through increased resonance (Albrecht & O'Brien, 1993; Myers & O'Brien, 1998), we expected to see an additional N400 amplitude reduction for the backward association condition relative to the forward association condition.

Additionally, there were a number of time points of interest, common in the psycholinguistic literature, at which the association conditions may differ. Semantic effects have been found during online linguistic processing as early as ~160 ms (Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006). Linguistically evoked responses occurring later include those reflecting high-level visual-semantic

Table 1. Sample passages for each experimental condition.

Text condition	Sample passage
FA	When the bear was awoke by the wandering chipmunk, he was filled with <u>rage</u>. The **<u>anger</u>** ruined her beautiful sweater.
BA	When the bear was awoke by the wandering chipmunk, he was filled with <u>anger</u>. The **<u>rage</u>** ruined her beautiful sweater.
Baseline #1	The bear was wandering early when he woke up the chipmunk. The **<u>anger</u>** he had experienced in early spring resembled the chipmunks today.
Baseline #2	The bear was wandering early when he woke up the chipmunk. The **<u>rage</u>** he had experienced in early spring resembled the chipmunks today.

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.

processing (N2; Dien, Frishkoff, Cerbone, & Tucker, 2003; Martin-Loeches, Hinojosa, Gomez-Jarabo, & Rubia, 1999) and attention (P2; Luck & Hillyard, 1994), stimulus classification and memory operations (P300, P600/late positive complex (LPC); Donchin, 1981; Donchin & Coles, 1988; Kutas, McCarthy, & Donchin, 1977; Rugg & Curran, 2007), as well as lexical-semantic processing (N400; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). We performed a principal component analyses for a data-driven approach to fractionating potential effects in the ERP waveforms.

In addition to the TC task, participants completed a word meaning judgement (MJ) task. The stimuli for this task were also asymmetrically associated words. Participants made meaning decisions to pairs that were strongly FA, strongly BA, 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 lexical-semantic processing within sparse and rich contexts. As different patterns of results have been seen across tasks such as lexical decision (Koriat, 1981) vs. naming (Seidenberg et al., 1984), responses in MJs may clarify our understanding of backward priming effects. In terms of behavioural (reaction time and/or accuracy) and neural (N400) responses on the MJ task, consistent with both previous lexical decision tasks (Chwilla et al., 1998; Koriat, 1981) and semantic judgement tasks (Kandhadai & Federmeier, 2010), we expected to find priming effects for both types of associated pairs relative to unrelated pairs.

Finally, because WTI effects may be linked to differences in reading comprehension skill (Yang, Perfetti, & Schmalhofer, 2005; Yang et al., 2007), we also had offline measures of reading skill (described in the methods section) that could be correlated with the experimental measures.

Methods

Participants

Thirty-one participants were recruited from the University of Pittsburgh student and staff community. All were right-handed, native English speakers between the ages of 18 and 35 years, with normal or correct-to-normal vision, without any history of head injury or epilepsy. Some participants were recruited from the Pittsburgh Adult Reading Database, which includes reading-related assessments including 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. The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), which includes sub-tests measuring word reading efficiency and non-word decoding, 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.

Materials

Word pairs were chosen such that their association strength was asymmetrical, that is, strong in one direction and weak or nonexistent in the other, according to the South Florida Association Norms (Nelson, McEvoy, & Schreiber, 1998). The association strength in the strong direction was at least .20, and in the weak direction no pair had association strength greater than .05 (Frishkoff, 2007). For the TC task, the 90 word pairs had mean association strength of .354 (SD = .14) in the strong direction and .017 (SD = .01) in the weak direction. For the MJ task, the 120 word pairs had mean association strength of .348 (SD = .13) in the strong direction and .014 (SD = .02) in the weak condition. The pairs did not differ across tasks in frequency or length ($ps > .5$). For the TC task, the constraint of choosing pairs that fit into the contexts necessarily led to frequency differences between words within a pair (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), $p < .001$). Because word pairs for the MJ task were chosen to match those in the sentence comprehension task, word in a pair also differed in 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), $p < .001$). Because the word pairs were seen in one or the other order approximately equal times across participants, effects of length, and frequency differences should have been attenuated.

A total of 90 two-sentence experimental texts were created (Supplementary table A). The first sentence of each passage contained one member of a pair of asymmetrically associated words (the *antecedent*); the second word (and first content word) of the second sentence contained the other member of the pair (the *critical word*). The antecedent and the critical words were chosen so that the texts were coherent and the words co-referential in both directions. In the FA text condition, the strong association strength was from the antecedent word to the critical word. In the BA text condition, the strong association strength was from the critical word back to the antecedent word. A baseline text condition was created by removing the associated antecedents from the first sentences, as well as making slight changes to word order to maintain coherence. The baseline texts were meaningful and coherent texts, not anomalous. The semantic content of the baseline and experimental texts was compared by the document-to-document tool on the Colorado University Latent Semantic Analysis website (http://lsa.colorado.edu/; Landauer & Dumais, 1997), which revealed a mean pairwise similarity of .804 (SD = .15) between the conditions. In all, four versions of each passage were created – two experimental and two baseline (one for each word in a pair) – and each version was assigned to a separate list, with the lists used approximately equally across participants (Table 1). No participant saw more than one version of a given text.

The 120 word pairs for the word MJ task (Supplementary table B) formed four sets of 30 pairs each that were assigned to three different pair conditions: 30 FA pairs with the strong association direction from prime to target, 30 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 lists were used an approximately equal number of times across participants.

Design and procedure

The experiment took place in a sound-attenuated, electrically insulated booth. After being fitted with an electroencephalogram (EEG) net, Participants were seated in an adjustable chair approximately 60 cm from the centre of a 15-in (38.1 cm) cathode ray tube (CRT) display. Which ERP task was first was counterbalanced across participants. The TOWRE was administered in between the ERP tasks.

During the TC task, participants passively read two-sentence passages for comprehension. Sentences appeared one word at the time in the centre of a computer screen for 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 allow for sentence wrap-up effects (Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Each text was preceded by a fixation cross (+). Following a randomly distributed 1/3 of the trials, participants answered a true–false comprehension question based on the meaning of the passage. Half were "true" and half were "false". Responses were registered on a response box. The comprehension questions were to encourage participants to 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 TC portion of the experimental session occurred in three blocks of trials of approximately 15 minutes each to allow for breaks. Stimuli were presented in random orders. Three practice texts preceded the experimental trials.

During the MJ task, pairs of words were presented on the screen one at the time. Upon presentation of the second (i.e. critical) word, participants made a button-press for a meaning similarity judgement, that is, whether the word was related in meaning to the first (prime) word. The prime word was presented for 1000 ms, followed immediately by the critical word for 2000 ms. Each trial began with a centred fixation cross (+) for 450 ms, followed by a blank screen for a random duration between 75 and 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 exposure duration of the second word. ("No Response" in red). The MJ portion of the experiment occurred as three blocks (approximately four minutes each) of randomly ordered trials.

Between the two ERP tasks, after a short break, the TOWRE was administered in the booth with audio recorded for offline scoring. The TOWRE consists of two tests of verbal fluency and decoding. In the word reading efficiency sub-test participants orally read as many words as they could in 45 second from a sheet of paper consisting of four columns of words ($n = 104$). In the non-word decoding sub-test participants were asked to orally decode as many non-words as they could in 45 seconds from a sheet of paper consisting of three columns of non-words ($n = 63$).

After completing the final ERP task, participants recruited outside of the Database took the Nelson–Denny vocabulary and comprehension tests. Participants were asked to complete as many of the 100 questions as they could in 7.5 minutes. The Nelson–Denny comprehension test features 6 text passages followed by comprehension questions ($n = 36$), and participants were asked to complete as many as they could in 15 minutes.

Apparatus and ERP recordings

ERP recordings were made from a 128 electrode Geodesic sensor net (Tucker, 1993) with Ag/AgCl electrodes (Electrical Geodesics, Inc., Eugene, OR). During recording, all impedances were kept below 40 kΩ (Ferree, Luu, Russell, & Tucker, 2001). A vertex reference was used during the recording. The EEG signals were digitally sampled at a rate of 500 Hz, and hardware filtered during recording between 0.1 and 200 Hz. A 30 Hz low-pass finite impulse response filter was applied to the recorded EEGs. For both tasks, EEGs were segmented from 200 ms before to 700 ms after the onset of the critical words (900 ms segments). To keep the number of trials equal across conditions for the MJ task, half of the 60 Unrl trials were deleted by removing all even number trials for every participant. Thus, prior to artefact detection, each participant had 30 trials each of the FA, BA, and Unrl conditions. Artefact removal on trials and electrodes was carried out in the same manner as reported previously (Stafura & Perfetti, 2014; Supplementary data C). This resulted in the removal of data for two participants that had more than 10 bad trials per condition on both tasks. The TC data of two additional participants were removed for the same reason.

For the remaining datasets, an average of seven electrodes (5.4%) were removed; these were replaced using spherical spline interpolation (Ferree, 2006) and re-referenced to the average of the channels. 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 EP Toolkit v2.41 (Dien, 2010) for PCA analysis, or to SPSS 19.0 for statistical analyses.

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

Results

Descriptive data

Table 2 displays descriptive data, along with full and partial correlations among the individual difference measures. Both mean comprehension scores (25.24) and vocabulary scores (64.07) were about one standard deviation above the mean of 6328 participants in the Pittsburgh Adult Reading Database, comprehension mean = 20.86 (SD = 5.9) and vocabulary mean = 49.05 (SD = 15.6). Mean standardised scores on the Word Reading (106.2) and Decoding (103.6) sub-tests of the TOWRE were within the average range. The correlations among the reading related tasks show the expected patterns of shared variance. A notable high correlation of .57 between vocabulary and comprehension is present even after their correlations with decoding and word recognition are removed.

Text comprehension

Accuracy on the comprehension questions was above 85% across conditions, indicating that participants were attending to the texts.

ERP analysis

Our analysis strategy for the TC task had two parts. One, we explored N400 effects in the raw ERP data in order to provide a comparison with previous findings for WTI (Stafura & Perfetti, 2014; Yang et al., 2007). To examine N400 ERP responses, mean amplitudes from 300 to 500 ms after the onset of the critical word were averaged across a three electrode clusters (Figure 1): a left parietal cluster (centred on P3), a central cluster (centred on Cz), and a right parietal cluster (centred on P4). These clusters cover a broad central-parietal region where N400 effects are most clearly visible. Two, we performed a PCA, providing a data-driven approach to the ERP data. The PCA used the EP Toolkit v2.41 (Dien, 2010) in MATLAB 8.2 (The MathWorks Inc., 2013).

Analysis of mean amplitudes

First, in the 300–500 ms time window, we found a greater positivity elicited by critical words in the BA texts compared with the FA and Baseline texts over left parietal electrodes. Second, we found greater reductions in N400 negativities elicited by critical words in the two association conditions (FA and BA) compared with the Baseline texts over central electrodes.

These findings come from A 3 × 3 repeated-measures analysis of mean amplitude variance (ANOVA) within this time window with Condition (FA, BA, Baseline) and Cluster (P3, Cz, P4) as within-subject factors. The analysis showed main effects of Condition $F(2,52) = 4.736, p = .018$, $\eta_p^2 = .154$, Cluster; $F(2,52) = 12.287$, $p < .001$, $\eta_p^2 = .321$ and a Condition × Cluster interaction: $F(4,104) = 2.641$, $p = .049$, $\eta_p^2 = .092$. We tested the source of the interaction by testing Condition at each Electrode Cluster. The P3 cluster showed a condition effect ($F(2,52) = 9.224$, $p < .001$, $\eta_p^2 = .262$) that was due to a greater positivity (as opposed to a reduced negativity) for the BA condition

Table 2. Participant descriptive information.

Participants: $n = 29$ (female = 17)		Full and partial correlations				
Variable	Mean (SD)	Variable	ND Comp	ND Vocab	TOWRE WR	TOWRE Decoding
Age	21.97 (3.7)	ND Comp		.554**	.301	−.041
ND Comp[a]	25.24 (5.4)	ND Vocab	.572**		.156	.322~
ND Vocab[a]	64.07 (14.1)	TOWRE WR	.210	.060		−.167
TOWRE WR[b]	106.21 (10.1)	TOWRE Decoding	−.231	.418*	−.173	
TOWRE Decoding[b]	103.62 (11.1)					

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

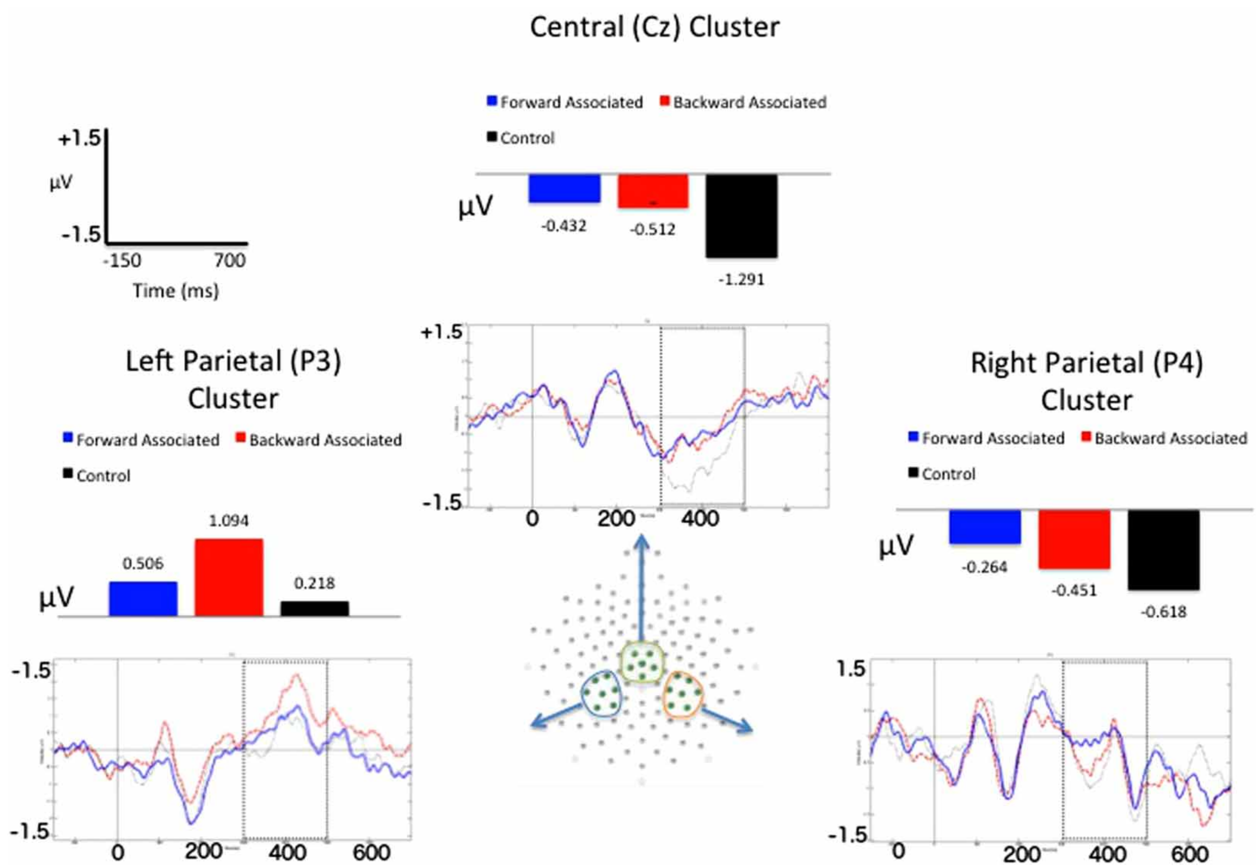

Figure 1. Waveforms and amplitude charts for the TC task ERP data. On the bottom is a schematic of the electrode net used in this study (anterior at the top), along with the three electrode clusters of interest (highlighted). On the right side of the schematic is the averaged waveform for the right (P4) 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 300–500 ms 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 300–500 ms time window for each condition, and for each cluster, is shown above each respective waveform. [To view this figure in colour, please see the online version of this Journal].

than either the FA or Baseline conditions (BA–FA: $t(26) = -2.806$, $p = .009$; BA–Baseline: $t(26) = 4.256$, $p < .001$). FA and Baseline conditions did not differ significantly (FA–Baseline: $t(26) = 1.385$, $p = .178$). These effects can be seen in Figure 1, leftmost panel.

Figure 1 also shows a Condition effect over the Cz cluster ($F(2,52) = 4.976$, $p = .018$, $\eta_p^2 = .161$). This reflects the greater reductions in the negative deflections for the FA and BA conditions relative to the Baseline condition (FA–Baseline: $t(26) = 2.494$, $p = .019$; BA–Baseline: $t(26) = 2.340$, $p = .027$), while FA and BA did not differ ($t(26) < 1$). The effect of Condition over the P4 cluster was not significant ($F(2,52) < 1$).

Post hoc *late positivity ERP analysis*

As can be seen in Figure 2, there was a striking left-lateralised positivity for the backward association condition throughout the 500–700 ms post-stimulus time window. A repeated-measures ANOVA of mean amplitudes over a broad left parietal-temporal region (electrodes E58/P7, E52/P3, E45/T7, and E36/C3) verified that this positivity was greater for the backward association condition than the forward condition: main effect of condition: $F(2, 52) = 4.217$, $p = .025$, $\eta_p^2 = .140$; FA–Baseline contrast: $t(26) = -1.024$, $p = .946$; BA–Baseline contrast: $t(26) = 1.679$, $p = .315$; FA–BA contrast, $t(26) = -2.998$, $p = .018$.

PCA analysis

To observe to more complete effects over the whole trial, we carried out a temporal PCA using a Promax rotation (Hendrickson & White, 1964) and the covariance matrix. (The rotation parameter was set to the default of 3, and Kaiser weighting was used.) A scree plot contrasting variance accounted for by each component against variance accounted for by random data suggested retention of 9 principal components (Horn, 1965) that accounted for 86.29% of the variance. Three components unrelated to experimental effects were discarded: An early component

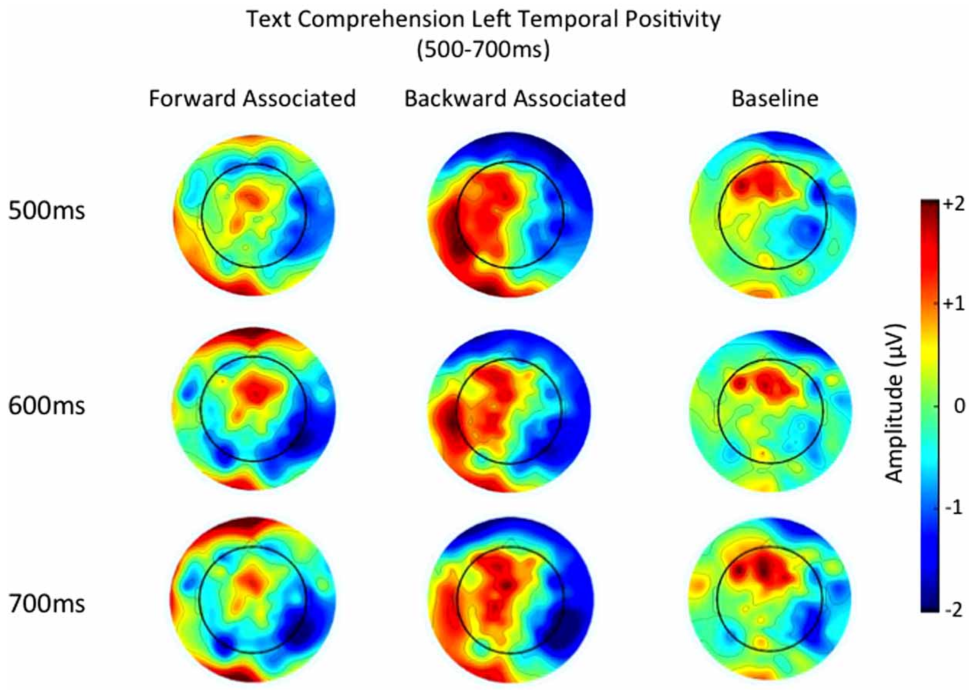

Figure 2. Topographic voltage maps for the TC conditions from 500–700 ms after stimulus onset. A left-lateralised positivity can be seen emerging around 500 ms after critical word onset for the FA and BA conditions; this positivity shows a peak at 500 ms for the FA condition, but extends throughout the time window for the BA condition. [To view this figure in colour, please see the online version of this Journal].

(peak at 28 ms) reflecting endogenous activity, a component peaking at 120 ms and corresponding to the P100 visual processing component, and a late component peaking at 596 ms that is typically found in ERP data, and assumed to reflect time-dependent trial-related drift. The remaining six principal components reflected co-varied voltage shifts that correspond to conventional ERP components defined by latency and polarity. We refer to the PCA components by using "F" for Factor, followed by the time in milliseconds of peak activity. These include the F160, F196, F284, F332, F380, and F424, which capture variance associated with the conventional N170, N2, P300, N400, and LPC ERP components, respectively (Figure 3).

For each factor, we examined the factor weightings at the electrode at which peak activity occurred, using a window centred on the time of peak activity. The time windows were 30 ms for the shorter early F160 and F196, and 50 ms for the remaining factors. For each principal component, we carried out a repeated-measures ANOVA with Condition as the within-subjects factor and a planned contrast of FA vs. BA. Finally, we used Bonferroni corrected contrasts to test the difference between each Association condition and Baseline. The results are reported below in order of temporal occurrence.

F160. This earliest factor, centred on electrode E60 (just posterior to P3) did not differ between conditions; $F(2, 52) = 2.091$, $p = .135$.

F196. The next factor distinguished reliably at a left parietal site (E58 or P7) between BA and Baseline, and marginally between BA and FA. Main effect of condition: $F(2, 52) = 3.797$, $p = .034$, $\eta_p^2 = .127$; FA–BA contrast, $t(26) = -1.953$, $p = .062$; BA–Baseline contrast, $t(26) = 2.934$, $p < .05$; FA–Baseline contrast, $t < 1$. Figure 3 illustrates the reduced negativity elicited on average by the BA texts.

F284. This factor, centred on electrode E90 (just right of O2), showed only a marginal condition effect: $F(2, 52) = 2.570$, $p = .089$. One can see a trend for a greater positivity in the Baseline relative to FA, with BA eliciting the least positivity (Figure 3). FA and BA did not differ $t(26) = 1.328$, $p = .196$.

F332. This mid-latency factor centred on electrode CZ distinguished between the two association conditions and the baseline condition $F(2, 52) = 3.685$, $p = .036$, $\eta_p^2 = .124$.: FA–Baseline contrast: $t(26) = 2.207$, $p < .05$; BA–Baseline contrast: $t(26) = 2.402$, $p < .05$; FA and BA did not differ $t(26) < 1$. This factor (Figure 3) appears to have captured the N400 effects seen over the central (Cz) cluster in the ERP amplitude analysis, where the same pattern of results was found (Figure 1).

F380. This mid-latency factor, centred on electrode E77 (just right of Pz), distinguished FA from both BA and baseline. Condition effect: $F(2, 52) = 7.541$, $p = .001$, $\eta_p^2 = .225$; FA–BA contrast, $t = 2.584$, $p = .016$; FA–Baseline contrast, $t = 3.647$, $p < .05$; BA–Baseline contrast, $t =$

1.439, $p = .486$. Figure 3 illustrates a reduced negativity elicited in the FA condition relative to the other conditions.

F424. The latest-occurring factor, centred on electrode E71 (Just anterior to O1) differed across all conditions, with BA eliciting the greatest positivity, followed by Baseline, and FA eliciting the smallest response (Figure 3). Main effect of condition: $F(2, 52) = 5.132$, $p = .01$, $\eta_p^2 = .165$; FA–BA contrast, $t = -3.484$, $p = .002$; FA–Baseline contrast, $t = -1.442$, $p < .05$; BA–Baseline contrast, $t = 1.469$, $p < .05$.

In summary, the temporal PCA analysis revealed early, middle and late components that distinguished between the association conditions. At the earliest time, F196 revealed a left hemisphere reduced negativity for the BA texts relative to baseline texts, and a reduced negativity for BA texts relative to FA texts ($p = .06$). A mid-latency component over central electrodes, F332, revealed a reduced negativity for both association conditions relative to the baseline condition, consistent with that seen in the ERP analysis over the central cluster (Figure 1). Later, F380 revealed a reduced negativity (mainly right hemisphere) for the FA texts relative to the other conditions. Unlike the F332 factor, there was no corresponding effect over the right parietal cluster in the ERP mean amplitude analysis. In the ERP data (Figure 1) the averaged waveforms diverged only in the first half of the a priori 300–500 ms window perhaps limiting mean differences over the full window. We note parenthetically that this highlights the advantage of PCA in revealing effects that are missed in traditional analyses. The latest component, F424, a positivity over left parietal electrodes distinguished between all conditions, with BA texts eliciting the greatest positivity and FA texts eliciting the smallest positivity, with baseline texts in between. This effect is consistent with the beginning of the positivity for the backward association condition in the *post hoc*, left temporal analysis (Figure 2).

Meaning judgements

Behavioural analysis

Table 3 shows the complete behavioural data (error rates and reaction times) for the MJ task. Participants judged both FA pairs and BA pairs as semantically related (96% and 95%, respectively), $t(28) = 1.00$, $p > .3$. For decision times, we compared FA and BA (related) trials with “Related” responses and Unrelated (Unrl) trials with “Unrelated” responses. Participants made “related” responses more quickly in the FA condition (588 ms) than the BA condition (627 ms). Unrelated trials took an average of 667 ms to make an “unrelated” response. T-tests revealed that all conditions differed reliably, all *p*s < .001.

ERP analysis

As with the TC trials, we examined the mean amplitudes from 300 to 500 ms from the onset of the critical (second) word in the MJ trials at the same three electrode clusters as used in the TC analysis (Figure 4): a left parietal cluster (centred on P3), a central cluster (centred on Cz), and a right parietal cluster (centred on P4).

Analysis of mean amplitudes

Two main results emerged. First was more positivity in the 300–500 window for the FA pairs relative to those in BA pairs and unrelated pairs. This difference extended over both right and left parietal electrodes. Second was a reduced negativity for the associated pairs (forward and backward) relative to those in unrelated pairs over central electrodes.

These results were confirmed by a 3 × 3 repeated-measures analysis of variance (ANOVA) on mean amplitudes in the 300–500 ms time window. Both Condition (FA, BA, Unrl) and Cluster (P3, Cz, P4) and their interaction showed significant effects: Condition: $F(2,56) = 6.768$, $p = .007$, $\eta_p^2 = .195$; Cluster; $F(2,56) = 27.853$, $p < .001$, $\eta_p^2 = .499$; Condition × Cluster $F(4,104) = 4.175$, $p = .007$, $\eta_p^2 = .130$. To examine the source of the interaction we tested Condition at each Electrode Cluster. Over the P3, Condition showed a marginal effect ($F(2,56) =$, $p = .075$, $\eta_p^2 = .095$), reflecting greater positivity for FA compared with BA and Unrl conditions (FA–BA: $t(28) = 2.545$, $p = .017$; FA–Unrl: $t(28) = 2.331$, $p = .027$). BA and Unrl were not different ($t(28) < 1$). These effects can be seen on the leftmost chart and waveform in Figure 4. Over the Cz cluster Condition was also significant ($F(2,56) = 7.593$, $p = .002$, $\eta_p^2 = .213$) due to reduced negativities for the FA and BA conditions relative to the Unrl condition (FA–Unrl: $t(28) = 3.179$, $p = .004$; BA–Unrl: $t(28) = 3.133$, $p = .004$). FA and BA did not differ ($t(28) < 1$). These effects can be seen in the centre chart and waveform in Figure 4. Finally, the P4 cluster also showed Condition effect ($F(2,56) = 7.318$, $p = .002$, $\eta_p^2 = .207$). FA produced less negativity than either BA or Unrl (FA–BA: $t(28) = 3.134$, $p = .004$;

Table 3. MJ mean (standard deviation) accuracy and reaction times.

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

Note: Accuracy refers to the proportion of “Related” responses to FA and BA word pairs, and “Unrelated” responses to unrelated pairs.

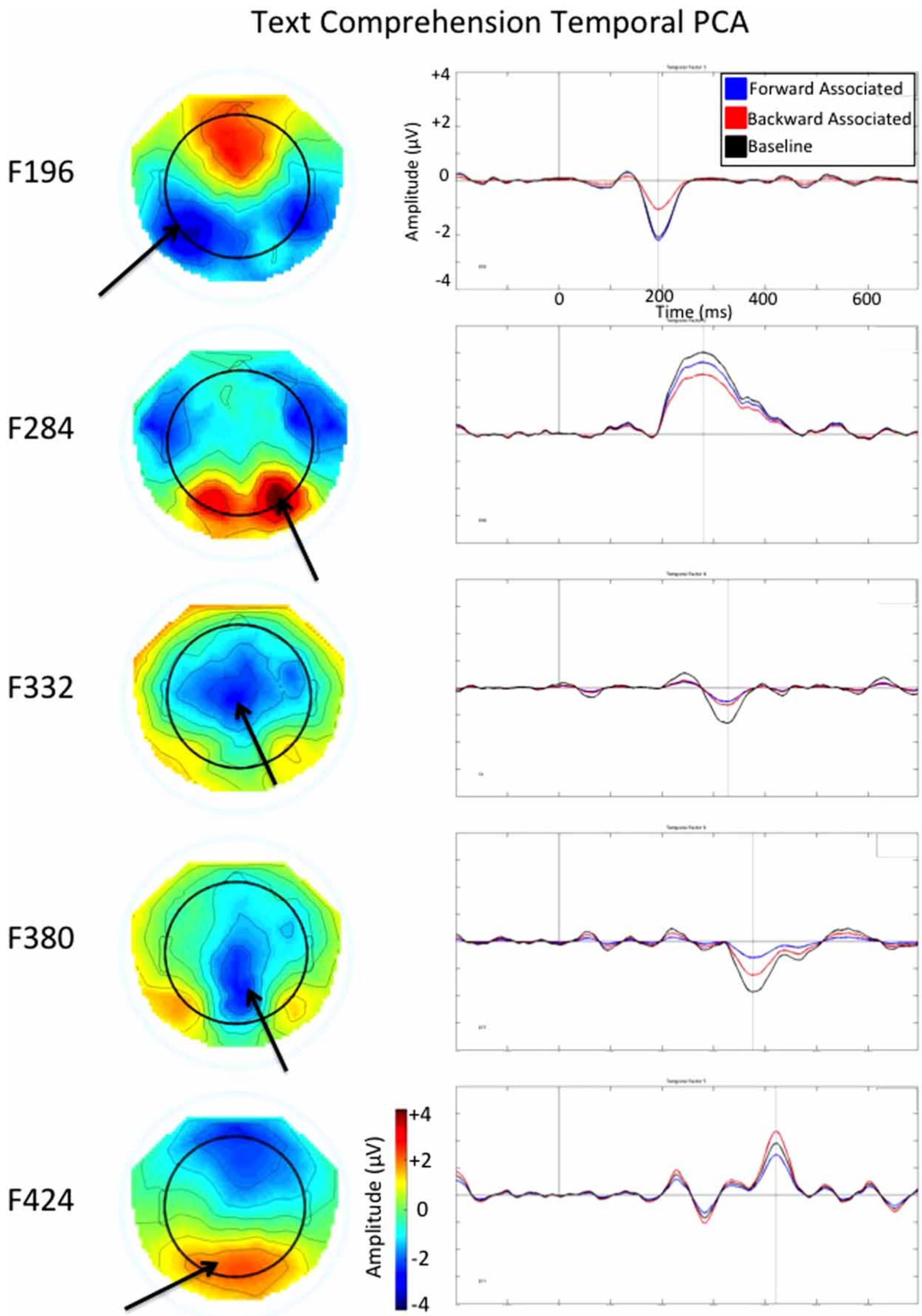

Figure 3. Temporal PCA components extracted from the TC data. Principal components are illustrated in temporal order from top to bottom. The far left column indicates the time of peak activity of each factor (e.g. F196 = 196 ms). The next column illustrates topographic voltage maps, averaged across all conditions. The black arrow indicates the site of peak activity for each topographic map. The far right column contains the grand average factor waveforms for each condition, scaled to millivolts, with positive plotted up (±4 µV). FA is in blue, BA is in red, and Baseline is in Black. [To view this figure in colour, please see the online version of this Journal].

FA–Unrl: $t(28) = 3.431$, $p = .002$). BA and Unrl conditions did no differ not differ, $t(28) = 1.175$, $p = .250$). These effects can be seen on the rightmost chart and waveform in Figure 4.

PCA analysis

The temporal PCA used the same parameters as that for the TC data. The scree plot suggested retention of 8 principal components, accounting for 92.35% of the variance of the ERP data. After discarding five components that were not sensitive to experimental conditions,[1] the remaining three principal components reflected co-varied voltage shifts that correspond to conventional ERP components and also to the factors extracted from the text data. Again, we refer to the PCA by "F" for Factor and the latency in ms to peak activity: F167, F224, and F324; these PCA components captured variance associated with the conventional N170, N2, and N400 ERP components, respectively (Figure 5).

For each factor, we examined the factor weightings at the electrode at which peak activity occurred, using a window centred on the time of peak activity: 30 ms for the short, early F168 and F224, and 50 ms for the F324. Each principal component was examined in a repeated-measures ANOVA with Condition as the within-subjects factor and further tested with planned contrasts of FA vs. BA and Bonferroni corrected contrasts between each Association condition and the Baseline condition. The results are reported below in order of temporal occurrence.

F168. This early factor, centred on electrode E65 (just anterior to O1) did not differ among conditions; $F(2, 56) = 1.115$, $p = .332$.

F224 centred on electrode E58 (P7) also did not differ among conditions; $F(2, 56) < 1$.

F324. This mid-latency factor centred on electrode E6 (just anterior to the vertex) distinguished the two association conditions from the unrelated condition, $F(2, 56) = 11.429$, $p < .001$, $\eta_p^2 = .290$; FA–Unrl contrast: $t(26) = 4.392$, $p < .05$; BA–Unrl contrast: $t(26) = 3.446$, $p < .05$; FA and BA did not differ, $t(28) = 1.306$, $p = .202$. This factor (Figure 5) seems to capture the N400 effects seen in the ERP amplitude analysis at the central electrodes (Figure 4), where the same pattern of results was found.

Correlations with individual difference measures

In order to examine associations between online task performance and offline comprehension ability, a composite language z-score ("comprehension skill") was created for each participant by averaging vocabulary and reading

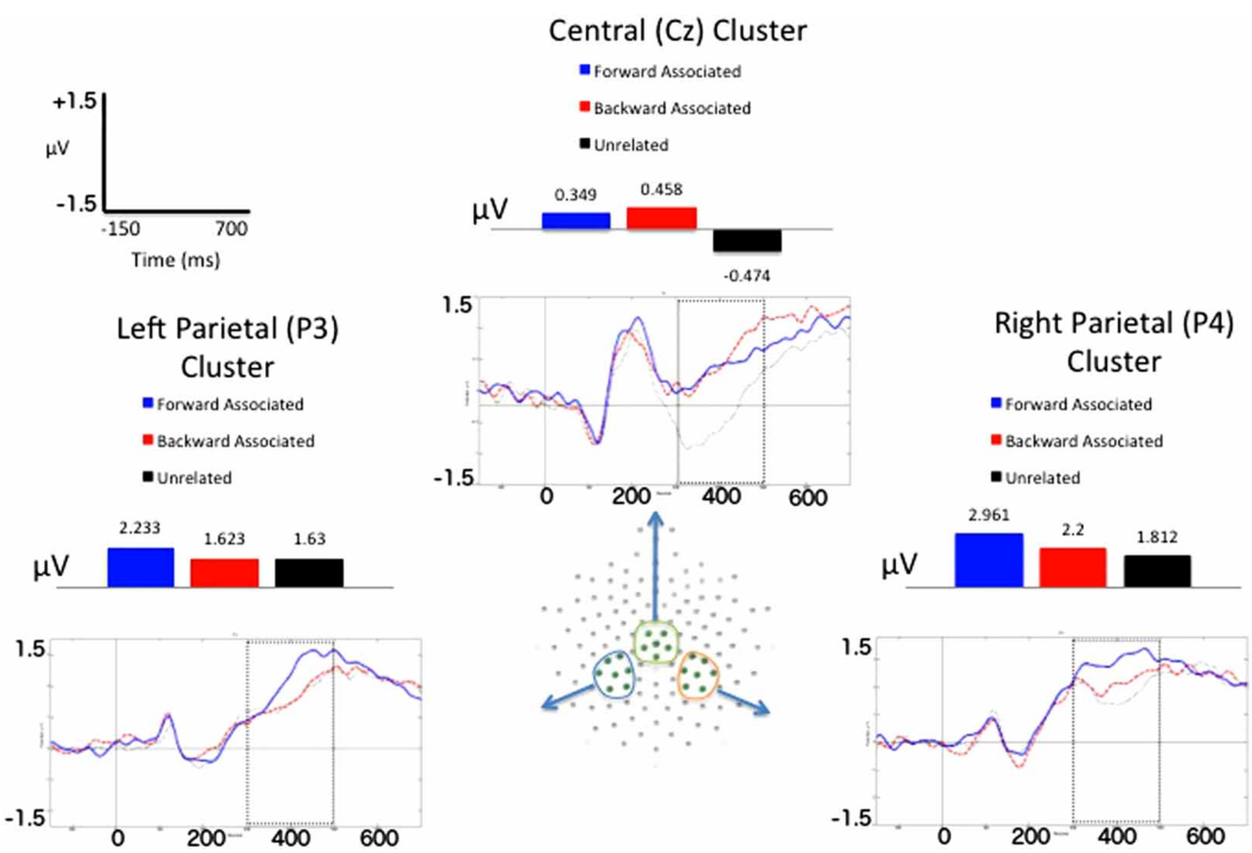

Figure 4. Waveforms and amplitude charts for the MJ task ERP data. On the bottom is a schematic of the electrode net used in this study (anterior at the top), along with the three electrode clusters of interest (highlighted). On the right side of the schematic is the averaged waveform for the right (P4) 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 300–500 ms 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 300–500 ms time window for each condition, and for each cluster, is shown above each respective waveform. [To view this figure in colour, please see the online version of this Journal].

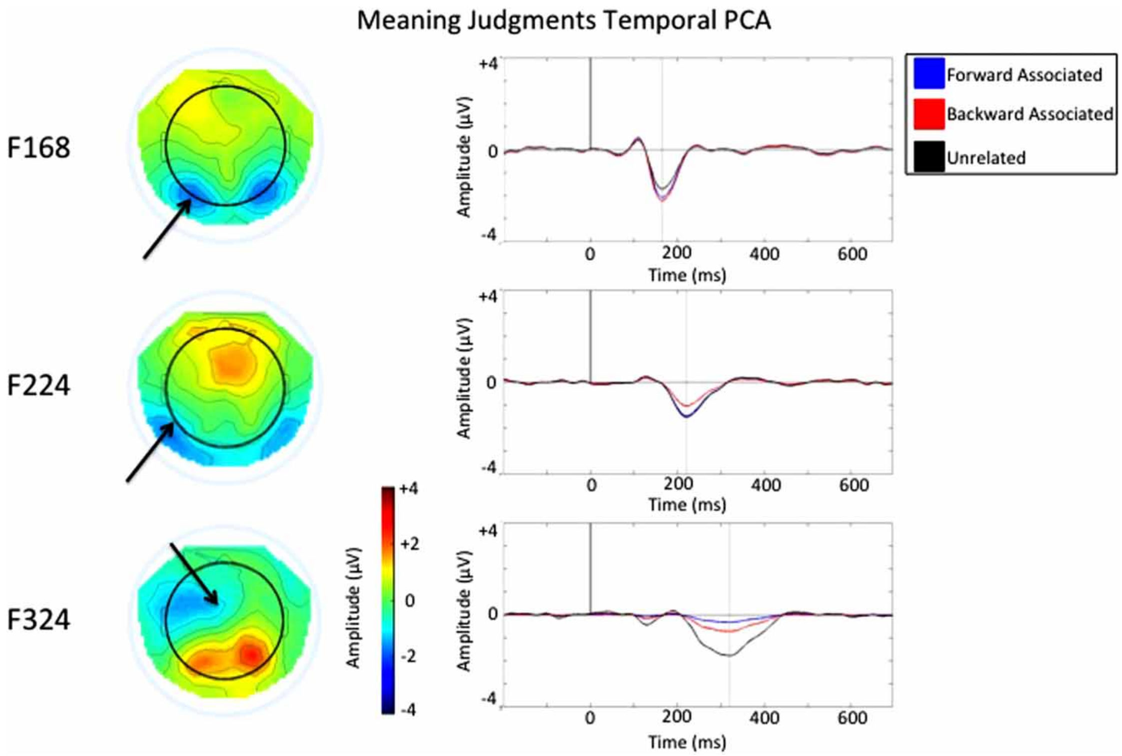

Figure 5. Temporal PCA components extracted from the word MJ data. Principal components are illustrated in temporal order from top to bottom. The far left column indicates the time of peak activity of each factor (e.g. F196 = 196 ms). The next column illustrates topographic voltage maps, averaged across all conditions. The black arrow indicates the site of peak activity for each topographic map. The far right column contains the grand average factor waveforms for each condition, scaled to millivolts, with positive plotted up (±4 µV). FA is in blue, BA is in red, and Unrelated is in black. [To view this figure in colour, please see the online version of this Journal].

comprehension z-scores. A composite score captures the shared variance between vocabulary and comprehension reflected in the full and partial correlation seen in Table 2.

MJ behavioural data. For MJ times, relatedness priming (FA and BA minus unrelated) was associated with composite z-scores ($r = .546$, $p = .002$), and forward association priming (FA-unrelated) was associated with the composite z-scores ($r = .492$, $p = .007$). The composite z-scores were not reliably associated with backward priming effects ($r = .302$, $p = .111$). Additionally, word reading ability was negatively correlated with the difference in response latency to FA and BA conditions; Sight Word Reading: FA–BA, $r = -.392$, $p = .035$.

TC ERP measures. The composite z-scores were negatively associated with N400 differences over the Cz cluster during the reading of critical words in BA texts relative to baseline texts; BA–Baseline, $r = -.400$ $p = .039$.

TC PCA measures. The composite z-scores were marginally associated with F324 differences between the BA and baseline conditions. BA–Baseline, $r = -.340$, $p = .083$. Because the F324 factor captures the N400, the ERP- and PCA-based skill correlations are convergent, although weaker in the PCA.

Discussion

The results show prospective and retrospective lexical association influences on WTI during the reading of connected texts. ERP records during the reading of a critical word at the beginning of a sentence expose a time-course of word processing that is sensitive to lexical relations to a word in the preceding sentence. First, the results show that a classic message-level effect in ERPs (the N400) is not only obtainable across a sentence boundary, but is indifferent at one time point and over one scalp location (Cz) to the direction of association. The across-sentence boundary effect replicates Yang et al. (2007) and the message-level effect adds to a related finding by Stafura and Perfetti (2014) that the cross-boundary effect is indifferent to the strength of forward word-association. However, the time-course of effects shows that the direction of association did matter in other ways: FA (sentence 1 word to sentence 2 word) and BA (sentence 2 word to sentence 1 word) effects are observable at different time points and at different scalp locations. In particular, the PCA extracted both an early (*F196*ms) and a later (*F434*ms) retrospective effect – distinguishing backward from forward association. An additional mid-latency prospective effect (*F380*ms) distinguished forward association from both baseline and backward association. Temporally in between these distinctly directional effects, both forward and backward associations showed an influence at the F332 (corresponding to a central N400).

Before placing these results in the perspective of TC, we first discuss the results of our MJ task, which used similarly asymmetrically associated words. Consistent with previous findings in lexical decisions tasks (Chwilla et al., 1998), we found similar N400 reductions for FA and BA pairs relative to unrelated pairs over central electrodes. However, over more posterior electrodes, particularly over right parietal sites, FA pairs elicited reduced negativities relative to backward and unrelated pairs, consistent with previous findings in MJ tasks (Kandhadai & Federmeier, 2010). Additionally, meaning decision times favoured pairs that were FA over pairs that were BA. These results contrast with results reported for lexical decision tasks, where equivalent priming effects were found for forward and BA pairs (Koriat, 1981) and naming experiments that found no effect of backward association (Peterson & Simpson, 1989; Seidenberg et al., 1984). It is likely that association direction mattered in our experiment because our MJ task explicitly required access to lexical semantics, allowing semantic priming from the first to the second word to reduce the time to evaluate the second word's meaning. The correlation between behavioural forward priming and composite comprehension scores supports a higher-level reliance on meaning processing in this paradigm.

We return now to the TC results to place them in the context of comprehension research. We consider first the results related to the N400, which is the component observed in the cross-sentence WTI paradigm (Stafura & Perfetti, 2014; Yang et al., 2007) and in other comprehension paradigms (Ditman et al., 2007; Otten & Van Berkum, 2008; Van Berkum et al., 1999). In the present study, the N400 was affected by lexical associations between the critical word at the beginning of a second sentence and a co-referential antecedent in the first sentence. An N400 effect was seen over central electrodes both when the association direction was forward and when the direction was backward. These effects were captured in both grand average waveforms in the 300–500 ms window and in a mid-latency principal component (F332) maximal over central locations (Figures 1 and 3). Stafura and Perfetti (2014) interpreted their finding that the N400 effect in this window was not sensitive to forward association strength as evidence that the N400 reduction (relative to baseline) had its source at the message level, rather than the lexical level. Beyond this indifference to association strength and direction, however, are the specific directional effects that emerged at other time points.

We take backward association effects as evidence for memory-based message-level integration, in which the process is boosted by the association from the currently read word to a word in memory. N400 effects can reflect both forward anticipatory processes (Federmeier, 2007; Lau, Almeida, Hines, & Poeppel, 2009) and backward integration processes (Brown & Hagoort, 1993). Although the importance of predictive processes has been emphasised in studies that examine within-sentence effects, the backward association effect in our results suggests a strong role for memory processes as the reader begins a new sentence, where prediction of a particular word might be less helpful. Reading across-sentence boundaries involves a memory-based integration of a word meaning with text meaning. The memory processes include access to text elements in working memory and to text elements reactivated in longer-term memory (Ericsson & Kintsch, 1995). The resonance memory mechanism postulated by O'Brien, Plewes, and Albrecht (1990) for TC – a recurring memory activation process that has access to both a memory for the text and semantic memory – can be applied to memory-based WTI.

An additional perspective on the importance of memory processes (as opposed to predictive processes) in reading is the cognitive load of TC. First, recent ERP evidence shows that predictive processing is reduced in comprehension situations that limit time for strategic processing (Wlotko & Federmeier, 2015). Second, backward priming is less sensitive to memory load than forward priming (Heyman, Van Rensbergen, Storms, Hutchinson, & De Deyne, 2015), and may compensate in difficult reading conditions (Thomas, Neely, & O'Connor, 2012). If backward priming partly mimics memory-triggering processes in reading, this implies memory processes actually serve the management of the reader's cognitive load. In our BA text condition, access to an associated word in the previous sentence can represent, more generally, the accessibility of previously read text elements (words, referents, or propositions) during comprehension. This implies that reading words triggers access to prior text elements through memory and that this process continually tunes the mental representation of the text. The read word becomes an automatic retrieval cue for elements of the mental model that are refined by the contribution of the read word (Albrecht & O'Brien, 1993; O'Brien et al., 1990; Perfetti & Stafura, 2014; Ratcliff, 1978; Ratcliff & McKoon, 1988). We add the caution that our current evidence does not rule out that the retrieval involves a word or a proposition (text model elements) rather than a referential component of a mental model.

Beyond an N400 that reflects general message-level integration, we found evidence for lexical association effects in a time period that overlapped the central N400. This principal component, which was maximal at right parietal sites, peaked ~50 ms after the central N400 component and showed a reduced negativity for the FA texts relative to other texts (Figure 3). The fact that backward association produced a smaller and less reliable effect at the right parietal site suggests that this component is sensitive to both lexical-level and message-level information, but especially to the lexical level.

Two previous studies provide a comparison to this second mid-latency negativity; in these studies, a right parietal negativity within a time window similar to the classic N400 was sensitive to retrospective lexical processes (N400RP; Dien et al., 2006; Franklin, Dien, Neely, Huber, & Waterson, 2007). Franklin et al. recorded EEG while participants performed primed lexical decisions, with symmetrically associated, asymmetrically associated, and unrelated prime-target pairs. At long SOAs, BA targets elicited reduced negativities relative to unrelated targets, but only over right parietal regions, not more central regions where symmetrical and FA targets elicited reduced N400 responses. The authors interpret these results as demonstrating parallel processes of expectancy and retrospective semantic matching, with the former indexed by central scalp negativities (N400), and the latter indexed by right parietal negativities (N400RP). Our results differ, in that BA texts elicited similar N400 reductions as FA pairs over central scalp locations, and smaller N400 reductions relative to FA pairs over right parietal regions. Although the central effects could perhaps be explained by message-level expectancy processes, it is not clear why the FA texts would show a larger right parietal N400 effect than the BA texts if this component is particularly sensitive to retrospective matching processes (Franklin et al., 2007). There were a substantial number of differences between items and paradigms between studies, not least being the difference between lexical decisions and passive reading of meaningful text. So, although the overlap in findings of a right parietal negativity is intriguing, more research is needed to discover to functional significance of this component.

Thus, the two N400s identified here are responsive to different aspects of WTI. That multiple processes contribute to the temporally and spatially extended N400 component is not surprising, given the complexity of the semantic system (e.g. Binder, Desai, Graves, & Conant, 2009). And, even though localisation with scalp EEG is indeterminate, different patterns of activation over different clusters of electrodes is suggestive of differing neural generators, or shifting weightings on distributed generators. The present results suggest that the existence of a co-referential antecedent is sufficient to facilitate integration of a given word, but prospective lexical association gives an additional facilitative boost to the WTI process, at a slightly later time point.

Backward association, however, showed an earlier effect. The temporal PCA extracted a factor maximal over left temporal-parietal electrodes that peaked around 200 ms after word onset. This component reflected a reduced negativity when the word was associated to an antecedent word in the preceding sentence in the backward direction (critical word to antecedent word) compared with both baseline and FA conditions (Figure 3). Semantic effects around 200 ms have also been reported by Dien and colleagues (Dien et al., 2003; Dien & O'Hare, 2008). This early PCA component may be related to the N200 lexical recognition component (Martin-Loeches et al., 1999). In our text backward association condition, recognition of the critical word may have benefited from meaning features of the antecedent word shared by the critical word, which it has reactivated. In the FA texts, the critical word is not able to reactivate these features sufficiently to offset their decay across text and time. The lack of a significant association effect in a MJ PCA factor at around the same time and location (F224) as this text factor suggests a distinction between text reading meaning processes and simple meaning matching processes.

Backward association also produced a later effect in a PCA component that reflected a positive-going deflection over parietal electrodes, maximal around 425 ms after word onset (Figure 3). This component separated all text conditions, with backward texts the most positive, followed by baseline texts, then forward texts. This posterior positivity may correspond to left parietal positivities associated with memory operations (left parietal Old/New effect; P600), including those involved in recognition memory (Rugg & Curran, 2007; Rugg et al., 1998) and in updating discourse representations (Burkhardt, 2007). In the present experiment, this may reflect a resonant memory process that is evoked by the critical word. Upon encountering this word in the second sentence, passive resonance processes (O'Brien, Rizzella, Albrecht, & Halleran, 1998) lead to a co-activation of the co-referential information in the reader's situation model and the meaning of the critical word. While this co-activation can occur in all conditions, it is greater in the backward condition where the encoding of the critical word activates meaning features that are associated with the antecedent word concurrently in memory. Additionally, by acting as a retrieval cue, the critical word may lead to the retrieval of the propositional structure constructed from the preceding sentence (Ericsson & Kintsch, 1995; O'Brien et al., 1990; Ratcliff & McKoon, 1988), leading to a long-lasting positivity for the backward association texts (Figure 2).

Although the memory-based account of the late positivity (Rugg et al., 1998) provides a good fit to the memory-based theoretical framing of the integration process, we note that a similar P600 component is associated with syntactic violations (Osterhout & Holcomb, 1992), semantic implausibilities (Kuperberg, 2007) and the comprehension of metaphors (Coulson & Van Petten, 2002). However, the memory-based interpretation seems more consistent with our text manipulations, which involved no syntactic or semantic irregularities. In this respect, it is interesting that in our results, the baseline condition elicited a greater positivity than forward texts. This is consistent with the assumption that the baseline condition requires the reader to establish a new discourse structure in memory (Burkhardt, 2007; Gernsbacher, 1990, 1997). The greatest positivity was found for backward texts, which may result from extended processing of the previous antecedent(/proposition) due to semantic matching driven by the critical word. A possible outcome of this would be enhanced memory for the re-processed text (O'Brien & Albrecht, 1991; O'Brien et al., 1990).

Together, the ERP and PCA results of the TC task reveal a series of partially overlapping indicators of different aspects of the WTI process. Within 200 ms, the orthographic recognition process leads to lexical activation that spreads to semantically associated words and, through a rapid interaction of word meaning and context (e.g. Kintsch & Mross, 1985), brings about the selection of context-relevant meaning, This is the initial phase of the essential resonant memory process, which is enhanced by "backward" association from (currently read) word to text (memory). From 300 ms to about 500 ms, two overlapping aspects of an N400 reflect the increasing integration of word-to text, an essentially message-level effect (i.e. controlled by comprehension of text propositions), but one that is also facilitated by lexical-semantic association that activates meaning features that turn out to be related to the critical word; thus, the expectancy related N400. Integration processes continue through more explicit memory updating reflected in a posterior positivity signalling modification of the events represented in the mental model. This account inserts memory resonance into explanations of integration processes, asserting not merely that such processes involve the message level, but also providing a hypothesis about a mechanism that brings them about.

An additional intriguing set of results was that backward texts elicited greater positivity (or less negativity) than other conditions over left parietal sites during reading (Figure 1), and FA word pairs elicited less negativity (or greater positivity) than other conditions over right parietal sites during MJs (Figure 4). Although task differences are involved, it is interesting to consider these hemisphere differences in light of models of lateralised language processing. Multiple models predict a greater role for the left hemisphere in prediction and expectancy, and a greater role for the right hemisphere in retrospection/maintenance and integration (Dien, 2008; Federmeier, 2007). According to these models, the right lateralised effect over right parietal sites would have been generated by left temporal structures that contribute prominently to the classic N400 response (Halgren et al., 2002; Van Petten & Luka, 2006), which is highly sensitive to expectancy (Kutas & Federmeier, 2011; Lau et al., 2009). The left lateralised positivity, then, could have been generated by right hemisphere areas involved in integrative discourse processing (Bookheimer, 2002; St George, Kutas, Martinez, & Sereno, 1999). Of course, these are possibilities without direct evidence in the present study. (For discussion of relevant models see Dien, 2008; Federmeier, 2007; Long, Baynes, & Prat, 2005; Paivio, 1991)

Finally, we highlight the individual differences we observed in both TC and MJs. For the TC task, comprehension skill correlated negatively with the difference in N400 ERP amplitude between the backward association and baseline conditions. This suggests that highly skilled comprehenders receive less benefit from lexical associations in the WTI process. To put it another way, highly skilled comprehenders respond to the baseline condition in a way that is less differentiated from a co-referential condition. Highly skilled comprehenders also showed a reduced influence of association in the MJ task, where the difference in RTs between related and unrelated pairs was negatively correlated with both comprehension skill and vocabulary knowledge. Very skilled comprehenders are supported by their knowledge of word meanings and their general comprehension skill and less dependent on priming by context, a conclusion with parallels in research on skill differences among children in their dependence on context for word identification (Perfetti, Goldman, & Hogaboam, 1979; West, Stanovich, Feeman, & Cunningham, 1983).

The results of this study extend the results of previous studies on WTI and provide a plausible account of integration effects. While the N400 ERP results replicate the findings of a paraphrase effect (Stafura & Perfetti, 2014; Yang et al., 2007), the additional components identified through PCA revealed an interactive, extended timecourse of neural responses. The various ERP indicators reflect processing at multiple levels – message understanding, lexical association, and text memory – that occur during word processing and integration. A more specific and important conclusion comes from the demonstration that memory resonance is part of the explanation for integration processes. This is not so much an addition to message-level and lexical-level explanations, but rather a conclusion that memory resonance initiated by reading a word is the mechanism by which the message level asserts its effect.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported in part by the NICHD [grant number R01HD058566-02] to Charles A. Perfetti.

Supplemental data

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/23273798.2015.1062119.

Note

1. Discarded were a pre-baseline component reflecting processing of the prime word, a component peaking at 120 ms that reflects early visual processing captured by the P100 ERP component, a component peaking at 384 ms that was essentially a sinusoidal wave, and two late components peaking at 560 and 696 ms that reflected response-related activity and artefactual/time drift across trials, respectively.

ORCID

Joseph Z. Stafura http://orcid.org/0000-0001-5616-6913

References

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 1061–1070. doi:10.1037/0278-7393.19.5.1061

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264. doi:10.1016/S0010-0277(99)00059-1

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. doi:10.1093/cercor/bhp055

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. doi:10.1006/jmla.1995.1034

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*(1), 151–188. doi:10.1146/annurev.neuro.25.112701.142946

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*(1), 34–44. doi:10.1162/jocn.1993.5.1.34

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. doi:10.3758/BRM.41.4.977

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Neuroreport*, *18*(17), 1851–1854. doi:10.1097/WNR.0b013e3282f1a999

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. doi:10.1016/j.jml.2006.07.005

Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition*, *14*, 509–522. doi:10.3758/BF03202522

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. doi:10.1080/713755773

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. doi:10.1037/0278-7393.31.1.129

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, *30*, 958–968. doi:10.3758/BF03195780

Dien, J. (2008). Looking both ways through time: The Janus model of lateralized cognition. *Brain and Cognition*, *67*(3), 292–323. doi:10.1016/j.bandc.2008.02.007

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*(1), 138–145. doi:10.1016/j.jneumeth.2009.12.009

Dien, J., Franklin, M. S., & May, C. J. (2006). Is "blank" a suitable neutral prime for event related potential experiments? *Brain and Language*, *97*(1), 91–101. doi:10.1016/j.bandl.2005.08.002

Dien, J., & Frishkoff, G. A. (2005). Principal components analysis of event-related potential datasets. In T. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 189–208). Cambridge: MIT Press.

Dien, J., Frishkoff, G. A., Cerbone, A., & Tucker, D. M. (2003). Parametric analysis of event related potentials in semantic comprehension: Evidence for parallel brain mechanisms. *Cognitive Brain Research*, *15*, 137–153. doi:10.1016/S0926-6410(02)00147-7

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. doi:10.1016/j.brainres.2008.09.045

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. *Language and Cognitive Processes*, *22*(6), 793–827. doi:10.1080/01690960601057126

Donchin, E. (1981). Surprise! … surprise? *Psychophysiology*, *18*(5), 493–513. doi:10.1111/j.1469-8986.1981.tb01815.x

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374. doi:10.1017/S0140525X00058027

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. doi:10.1037/0033-295X.102.2.211

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505. doi:10.1111/j.1469-8986.2007.00531.x

Ferree, T. C. (2006). Spherical splines and average referencing in scalp electroencephalography. *Brain Topography*, *19*(1–2), 43–52. doi:10.1007/s10548-006-0011-0

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. doi:10.1016/S1388-2457(00)00533-2

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Copper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27–85). Cambridge: MIT Press.

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*(5), 1053–1068.

Frishkoff, G. A. (2007). Hemispheric differences in strong versus weak semantic priming: Evidence from event-related brain

potentials. *Brain and Language*, *100*, 23–43. doi:10.1016/j.bandl.2006.06.117
Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*, 265–304. doi:10.1080/01638539709544994
Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage*, *17*(3), 1101–1116. doi:10.1006/nimg.2002.1268
Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30*, 1383–1400. doi:10.1016/j.neuroimage.2005.11.048
Hendrickson, A. E. & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology*, *17*, 65–70. doi:10.1111/j.2044-8317.1964.tb00244.x
Heyman, T., Van Rensbergen, B., Storms, G., Hutchison, K. A., & De Deyne, S. (2015). The influence of working memory load on semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 911–920. doi:10.1037/xlm0000050
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. doi:10.1016/j.cogbrainres.2003.10.022
Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. doi:10.1007/BF02289447
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. doi:10.1098/rstb.1981.0145
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354. doi:10.1037/0033-295X.87.4.329
Kandhadai, P., & Federmeier, K. D. (2010). Hemispheric differences in the recruitment of semantic processing mechanisms. *Neuropsychologia*, *48*(13), 3772–3781. doi:10.1016/j.neuropsychologia.2010.07.018
Kintsch, W., & Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language*, *24*(3), 336–349. doi:10.1016/0749-596X(85)90032-4
Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition*, *9*, 587–598. doi:10.3758/BF03202353
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, Special Issue: Mysteries of Meaning*, *1146*, 23–49. doi:10.1016/j.brainres.2006.12.063
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. doi:10.1146/annurev.psych.093008.131123
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205. doi:10.1126/science.7350657
Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*, 792–795. doi:10.1126/science.887923
Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology*, *69*, 218–233. doi:10.1016/0013-4694(88)90131-9
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. doi:10.1037/0033-295X.104.2.211
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. doi:10.1016/j.bandl.2009.08.007
Ledoux, K., Camblin, C. C., Swaab, T. Y., & Gordon, P. C. (2006). Reading words in discourse: The modulation of lexical priming effects by message-level context. *Behavioral and Cognitive Neuroscience Reviews*, *5*(3), 107–127. doi:10.1177/1534582306289573
Long, D. L., Baynes, K., & Prat, C. S. (2005). The propositional structure of discourse in the two cerebral hemispheres. *Brain and language*, *95*(3), 383–394. doi:10.1016/j.bandl.2005.02.004
Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291–308. doi:10.1111/j.1469-8986.1994.tb02218.x
Martin-Loeches, M., Hinojosa, J. A., Gomez-Jarabo, G., & Rubia, F. J. (1999). The recognition potential: An ERP index of lexical access. *Brain & Language*, *70*, 364–384. doi:10.1006/brln.1999.2178
The MathWorks Inc. (2013). MATLAB (Version 8.2). Natick, MA: Author.
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. doi:10.1037/0278-7393.20.1.92
Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2), 131–157. doi:10.1080/01638539809545042
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, *11*, 264–336.
Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. *Psychology of Learning and Motivation*, *24*, 207–248.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. Retrieved from http://www.usf.edu/FreeAssociation/
Nelson, M. J., & Denny, E. C. (1973). *The Nelson-Denny reading test*. Boston, MA: Houghton Mifflin.
O'Brien, E. J., & Albrecht, J. E. (1991). The role of context in accessing antecedents in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 94–102. doi:10.1037/0278-7393.17.1.94
O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 241–249. doi:10.1037/0278-7393.16.2.241

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*(5), 1200–1210.

Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806. doi:10.1016/0749-596X(92)90039-Z

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*, 464–496. doi:10.1080/01638530802356463

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*, 255–287. doi:10.1037/h0084295

Perfetti, C., Yang, C.-L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology*, *22*, 303–318. doi:10.1002/acp.1419

Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition*, *7*(4), 273–282. doi:10.3758/BF03197600

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37.

Perfetti, C. A., & Stafura, J. Z. (2015). Comprehending implicit meanings in text without making inferences. In E. J. Obrien, A. E. Cook, & R. F. Lorch Jr. (Eds.), *Inferences during reading* (pp. 1–18). Cambridge: Cambridge University Press.

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. doi:10.1037/0278-7393.15.6.1020

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408. doi:10.1037/0033-295X.85.2.59

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. doi:10.1016/j.biopsycho.2008.05.002

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*, SI21–SI49. doi:10.1080/01690968908406362

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. doi:10.1016/j.tics.2007.04.004

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595–598. doi:10.1038/33396

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre-and post-lexical loci of contextual effects on word recognition. *Memory & Cognition*, *12*(4), 315–328. doi:10.3758/BF03198291

Stafura, J. Z., & Perfetti, C. A. (2014). Word-to-text integration: Message-level and lexical level influences in ERPs. *Neuropsychologia*, *64*, 41–53. doi:10.1016/j.neuropsychologia.2014.09.012

St George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain*, *122*(7), 1317–1325. doi:10.1093/brain/122.7.1317

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language*, *66*(4), 623–643. doi:10.1016/j.jml.2012.02.002

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. doi:10.1023/A:1026416225168

Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, *87*, 154–163. doi:10.1016/0013-4694(93)90121-B

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. doi:10.1016/S0022-5371(77)80027-3

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182. doi:10.1006/jmla.1999.2641

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. doi:10.1080/01690969308407586

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and language*, *97*(3), 279–293. doi:10.1016/j.bandl.2005.11.003

West, R. F., Stanovich, K. E., Feeman, D. J., & Cunningham, A. E. (1983). The effect of sentence context on word recognition in second-and sixth-grade children. *Reading Research Quarterly*, 6–15. doi:10.2307/747333

Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex*, *68*, 20–32. doi:10.1016/j.cortex.2015.03.014

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy*, *8*(2), 233–257. doi:10.1075/wll.8.2.10yan

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. doi:10.1037/0278-7393.33.1.55