NIH Public Access
Author Manuscript

Cognition. Author manuscript; available in PMC 2016 March 01.

Published in final edited form as:
Cognition. 2015 March ; 136: 135–149. doi:10.1016/j.cognition.2014.10.017.

Effects of Prediction and Contextual Support on Lexical Processing: Prediction takes Precedence

Trevor Brothers, **Tamara Y. Swaab**, and **Matthew J. Traxler**
University of California, Davis, Department of Psychology and Center for Mind and Brain

Abstract

Readers may use contextual information to anticipate and pre-activate specific lexical items during reading. However, prior studies have not clearly dissociated the effects of accurate lexical prediction from other forms of contextual facilitation such as plausibility or semantic priming. In this study, we measured electrophysiological responses to predicted and unpredicted target words in passages providing varying levels of contextual support. This method was used to isolate the neural effects of prediction from other potential contextual influences on lexical processing. While both prediction and discourse context influenced ERP amplitudes within the time range of the N400, the effects of prediction occurred much more rapidly, preceding contextual facilitation by approximately 100ms. In addition, a frontal, post-N400 positivity (PNP) was modulated by both prediction accuracy and the overall plausibility of the preceding passage. These results suggest a unique temporal primacy for prediction in facilitating lexical access. They also suggest that the frontal PNP may index the costs of revising discourse representations following an incorrect lexical prediction.

Keywords

prediction; N250; N400; sentence processing; event-related potentials

1. Introduction

Contemporary approaches to language processing emphasize forward looking, anticipatory processes (Elman, 1990; Altmann & Mirković, 2009; Kuperberg, in press). Rather than viewing comprehenders as passive processors of incoming linguistic input, these approaches view comprehenders as active constructors of meaning. Part of this meaning construction activity entails using prior knowledge and experience to generate expectations about how a discourse will unfold in the near future. The outcome of the anticipatory process, its success or failure, is viewed as a major factor contributing to the processing load that "bottom-up"

Address Correspondence to: Matt Traxler, UC Davis, Department of Psychology, 1 Shields Avenue, Davis, CA 95616, mjtraxler@ucdavis.edu, 530 754 9468 (o) 530 752 2087 (fax).

perceptual information imposes in a given context. Anticipatory, likelihood-driven processes are believed to contribute to lexical, syntactic, and discourse processing in a variety of theoretical frameworks (e.g., Gibson et al., 2013; Levy, 2008; Hale, 2001; Jaeger & Snider, 2013). Despite this strong emphasis on prediction and anticipation, there is very little direct evidence that successful versus unsuccessful prediction *per se* influences language processes as they unfold. Instead, the chain of inference is usually less direct. Contexts are constructed that make successful prediction easier or more difficult, and differences in processing difficulty across these contexts are then attributed to successful prediction. The current study takes a different approach that allows more direct assessment of the consequences of successful prediction on lexical access.

During natural language processing, words are typically embedded in a discourse which provides a broader context for interpreting meaning. Discourse context has been shown to facilitate lexical retrieval as well as memory for congruent words (e.g., Bransford & Johnson, 1972). Most broadly, prior linguistic information can activate networks of related concepts or event schemas which can facilitate processing (Shank, 1975). In more highly constraining contexts, comprehenders can also make predictions about specific lexical items that are likely to appear in the upcoming discourse[1]. For example, when listening to the sentence *"I could tell he was mad by the tone of his …"* the upcoming word *'voice'* will be processed more quickly (Schwanenflugel & Shoben, 1985, Traxler & Foss, 2000) and with higher perceptual accuracy (Miller et al., 1951, Jordan & Thomas, 2002) than if this word were presented in a less predictive context. A number of studies have shown that lexical prediction may influence comprehension processes before the critical word itself has even appeared. For example, words embedded in highly predictable contexts are skipped more frequently during natural reading (Rayner et al., 2011). While listening to continuous speech, participants will strategically fixate objects in a visual scene that are likely to be mentioned in the near future (Kamide, Altmann, and Haywood, 2003).

Electrophysiological studies of language processing have shown that a word's cloze probability, which is defined as the proportion of participants producing this word during an offline sentence completion task, correlates with the amplitude of the N400 component during reading comprehension (Kutas & Hillyard, 1984). Modulation of the amplitude of the N400 has been linked both to the facilitation of lexical integration processes (van Berkum, Hagoort & Brown, 1999), as well as lexical access and semantic memory retrieval (Kutas & Federmeier, 2000). More recently, it has been suggested that amplitude of this component is attenuated when features of a word have been pre-activated by the preceding context (for a review see Swaab et al., 2012).

In constraining contexts, the generation of specific lexical predictions could influence multiple levels of word processing via the pre-activation of phonological, morphological or semantic properties of the predicted word. Evidence for pre-activation comes from a variety of electrophysiological studies using sentence or discourse contexts. In these paradigms a

[1]Following Van Petten & Luka, 2012, we define *prediction* as a reader or listener's expectation "that a particular word (lexical item) will appear in the near future" (p. 179). Unlike more general forms of semantic anticipation, lexical prediction entails the pre-activation of word-specific features, such as phonological information or grammatical gender.

critical *probe* word is introduced that is either congruent or incongruent with the upcoming, anticipated lexical item. For example, Delong and colleagues (2005) found it was easier for readers to process the function word *'an*' while anticipating a phonologically congruent continuation like *'airplane'* as opposed to a phonologically incongruent one like *'kite'*. Other studies using this technique have found evidence for the pre-activation of animacy and semantic category information, as well as grammatical features like gender (Szewczyk & Schriefers, 2013; Boudewyn, et al., submitted; Van Berkum, et al., 2005; Wicha, et al., 2004).

While accurate predictions appear to facilitate lexical access and integration, there is also evidence for costs when a lexical prediction is disconfirmed. Behaviorally, these costs can result in longer reaction times for unexpected words appearing in highly constraining contexts (Schwanenflugel & Shoben, 1985). In the electrophysiological literature, it has been hypothesized that the costs of incorrect prediction are reflected in a late, post-N400 positivity (PNP). This positivity, which appears over frontal and left hemisphere electrode sites, is larger for unpredictable words in context (DeLong, Urbach, Groppe & Kutas, 2011), and this relationship appears to be modulated by sentence constraint (Federmeier, Wlotko, Ochoa-Dewald & Kutas 2007; but see Thornhill & Van Petten 2012 for conflicting results). Additionally, this frontal component appears to be distinct from other, posterior positivities which occur in response to anomalous or ungrammatical sentence continuations (DeLong, Quante & Kutas, in press). While little is known about the exact mechanisms underlying this frontal positivity, it has been hypothesized to reflect either the detection or resolution of disconfirmed predictions (see Van Petten & Luka 2012 for a review).

While these studies provide some evidence for both costs and benefits of predictive processing, nonetheless, there is a methodological limitation across this set of published studies which remains unresolved. Specifically, prior studies have not provided direct evidence on a trial-by-trial basis whether participants actually predicted the target stimulus. As a result, these studies have been unable to fully dissociate the effects of *specific* lexical pre-activation, from other sources of contextual support (e.g. semantic association or discourse plausibility). While some of the previously mentioned studies have side-stepped this issue by looking for the consequences of prediction earlier in the sentence, these studies have been unable to evaluate the effects of lexical pre-activation on the predicted words themselves. To address this issue, the current experiment used a paradigm that isolates the effects of lexical prediction from other sources of contextual facilitation.

In the experiment reported here, participants read moderately constraining (50% cloze) two-sentence passages, while trying to actively predict the final word of each. After each passage was complete, participants responded by button press whether their prediction was correct. By separately averaging ERP trials for predicted and unpredicted targets, we isolated processing differences at the final critical word that were uniquely driven by prediction accuracy[2]. In addition to these 50% cloze passages (whose final words should appear equally often as predicted and unpredicted targets) we also constructed passages that

[2]For a similar technique used to isolate the effects of memory encoding on subsequent recall (the difference due to memory, or DM effect) see Sanquist et al., 1980; Paller, Kutas & Mayes, 1987.

rendered these same critical words highly unpredictable (<1% cloze). This control condition allowed us to also compare *unpredicted* target words in low-cloze contexts to *unpredicted* targets in more supportive, medium-cloze contexts. Any differential ERP activity between these two conditions should index the amount of semantic or discourse-level facilitation provided by the preceding context, independent of lexical prediction.

1.1 Hypotheses

With this paradigm we tested three main hypotheses. If the generation of specific lexical predictions is an important mechanism by which contextual constraint facilitates lexical access, then self-reported prediction accuracy should have a clear influence on N400 amplitudes at the final critical word. In contrast, if we see a large influence of cloze probability on the N400 but fail to observe an effect of prediction accuracy, this would call into question the importance of lexical anticipation during sentence processing.

Secondly, if we do observe prediction-related ERP differences, the relative time-course of this effect would provide important constraints on models of lexical processing. Specific lexical predictions differ from other forms of contextual facilitation in that they entail the pre-activation of specific word forms. Unlike more general forms of semantic anticipation, lexical pre-activation may facilitate early stages of visual and orthographic processing (Lau, Holcomb & Kuperberg, 2013). To test this hypothesis, we investigated the effects of prediction accuracy and contextual constraint on the amplitude of the N250, an ERP component thought to reflect processing of visual word forms (Holcomb & Grainger, 2006).

Finally, this paradigm allowed us to more directly assess the processing costs which may be incurred following a disconfirmed lexical prediction. Based on current evidence (Van Petten & Luka, 2012), the frontal PNP may index: 1) the detection of prediction errors, 2) the inhibition of incorrectly pre-activated lexical items, or 3) a revision of the preceding discourse in light of new, unanticipated information. By separately examining the influence of prediction accuracy and contextual support on the PNP, we may have a better basis to determine which of these processing mechanisms it represents.

2. Materials and Method

2.1. Participants

Twenty-four undergraduates (16 females) from the University of California, Davis participated in the ERP study after providing informed consent. They were compensated with course credit. The mean age of participants was 20.6 years (range: 18–37, SD = 4.2). All were monolingual English speakers with no history of neurological or reading impairments. None had previously participated in any cloze testing or plausibility norming procedures for these experimental materials. All participants had normal or corrected-to-normal vision and were right-handed as determined by the Oldfield handedness inventory (1971).

2.2 Materials

The experimental stimuli consisted of 180 critical words, each appearing at the end of two, unique, two-sentence passages (360 passages in total, for examples see Appendix). The first passage of each pair was constructed to moderately constrain its final critical word. These medium-cloze passages were typically constrained toward two alternative completions that were equally likely given the preceding context. For example, in the passage: *"The author is writing another chapter about the fictional detective. To date, he thinks it will be his most popular novel.*", both the final critical word *'novel'* and an alternate completion *'book'* were predictable in context. The second passage of each pair was moderately constrained toward an unrelated completion, rendering the actual final word unpredictable yet semantically and grammatically acceptable. For example, in the passage: *"Everyone congratulated the chef on all his hard work. To date, he thinks it will be his most popular novel."* the final word *'novel'* would be unpredictable, yet still semantically coherent.

Two groups of 80, UC Davis undergraduates performed cloze norming on these passages. Participants saw only one version of each passage pair with the final word omitted, and they were asked to fill in the first word that came to mind. For the medium-cloze passages, the final critical words produced an average cloze probability of 50.7% (range 40–60%). In the low-cloze passages, these same critical words had a cloze probability of 0.9% (range 0–7%). The *constraint* of each passage was defined as the largest cloze-value obtained for any single word during the cloze norming study. By design, the low-cloze passages were all moderately constrained toward a particular target, but this target was never the final, critical word that was presented during the main experiment. Therefore, while the two experimental conditions differed significantly in cloze, $t(179) = 107$, $p < 0.001$, they did not differ in constraint (51.1% vs 51.8% for medium- and low-cloze respectively, $t < 1$).

As in the example shown above, the majority of passage pairs shared an identical second sentence (108 out of 180). For the remaining 72 pairs, the second sentences were always equal in word length, and the last four words, including the final critical word, were the same across the two versions (see Appendix for examples). These critical words had a mean SUBTLEX-US frequency of 109 per million, a mean length of 5.5 characters, and a mean concreteness rating of 496.

After the final 180 stimulus pairs were selected, a separate group of UCD undergraduates provided plausibility ratings for each complete passage. Two groups of 36 participants read equal proportions of medium- and low-cloze passages, one version from each pair, and were instructed to rate the plausibility of each using a 7-point scale (with 1 indicating the passage "makes perfect sense" and 7 indicating it "makes no sense at all"). As expected, the plausibility of the two passage types differed significantly (medium-cloze: mean = 1.83, SD = 0.50, low-cloze: mean = 4.34, SD=1.10, $t(179)=28.8$, $p < 0.001$).

The final 180 passage pairs were counterbalanced over three experimental lists, with each containing 120 medium-cloze and 60 low-cloze passages. This ensured that participants saw each context and critical word only once and that each target word appeared in equal proportions across the medium- and low-cloze passages.

2.3. Procedure

During EEG recording, participants were seated in an electrically-shielded, sound-attenuated booth with their heads resting comfortably on a chin rest. Stimuli were presented on an LCD monitor at a distance of 90cm. At the beginning of each trial, the first sentence of a passage appeared on the monitor, in full. The participants were instructed to read this sentence carefully for comprehension at their own pace. This sentence remained on the screen until the participant pressed a button, indicating they were ready to proceed with the passage. After this button press, a fixation cross appeared centrally for 1000ms to orient the participant's eyes. The second sentence of the passage was then presented, one word at a time, in the center of the screen. Each word appeared for 300ms with a stimulus onset asynchrony of 600ms. After the final critical word was presented, it was replaced by 1700ms of blank screen followed by a question mark.

For each trial, participants were instructed to read each passage carefully and to use the preceding context to try to predict the final word of the passage. After each passage was complete, participants were cued to respond "yes" or "no" via a button press to indicate whether they had successfully predicted the final word of the previous passage. Participants were told that there was no right or wrong answer on any trial, and that they should be as honest as possible in their responses. To minimize EEG artifacts, participants were instructed to withhold any blinks or button responses until after the appearance of the question mark. After a response was recorded, the experiment automatically proceeded to the next trial. After a short practice session, subjects were shown 6 blocks of experimental stimuli with a rest break between each. Both experimental list and response hand were counterbalanced across participants.

2.4. EEG Recording

The electroencephalogram (EEG) was recorded from 29 tin electrodes mounted in an elastic cap (Electro-Cap International; Eaton, OH). Additional electrodes were attached below, and to the side of each eye to monitor blinks and horizontal eye movements. All electrode impedances were kept below 5 kΩ. The EEG signal was amplified using a Synamps Model 8050 Amplifier (Compumedics Neuroscan) with a bandpass of 0.05Hz–30Hz. The signal was digitally recorded at a sampling rate of 250Hz.

After EEG recording, the data were re-referenced to an average of the left and right mastoids, and independent component analysis (ICA) was used to isolate and remove EEG components related to eye blinks. Single-trial waveforms were then screened for amplifier drift, muscle artifacts and eye movements, and any epochs containing these artifacts were rejected prior to analysis (1.2% of all trials). ERPs were calculated by averaging individual EEG epochs, time-locked to the presentation of each critical word. All ERP epochs were 1300ms in length including a 300ms pre-target baseline. Prior to off-line averaging, ERP trials were sorted based on sentence type (medium-cloze vs low-cloze) and prediction accuracy (predicted vs unpredicted). Except where indicated, statistical analyses were performed on individual subject ERP averages calculated in each condition. For all analyses with more than 1 degree of freedom, the reported p-values were first adjusted using the Greenhouse-Geisser correction.

3. Results

3.1. Behavioral results

In the medium-cloze condition, participants correctly predicted the final critical word for 48.1% of the passages (S.D. = 6.6%). This value was quite similar to the off-line cloze values previously obtained for these items: 50.7%. For the low-cloze condition, participants reported a correct prediction for 2.4% of the sentence-final words (S.D. = 2.2%), which was slightly higher than the off-line cloze value of 0.9%. Overall, prediction accuracy differed substantially between the medium-cloze and low-cloze passages, $t(23)=32.75$, $p < 0.001$, $d = 13.7$, suggesting participants understood the task and were attending carefully to the preceding context. As expected, only a small proportion of low-cloze items were reported as "predicted". Because of the small number of trials available per participant, this condition (low-cloze, predicted) was not included in the subsequent ERP analyses.

3.2. ERP results

Grand average ERP waveforms for each of the three critical conditions: predicted medium-cloze words, unpredicted medium-cloze words, and unpredicted low-cloze words (predicted, unpredicted, and low-cloze) are plotted in Figure 2. These conditions showed overlapping baseline activity followed by visual P1 and N1 components that were maximal at occipital electrode sites. Following these components, the waveform for predicted final words showed a broadly distributed positivity with a peak occurring at approximately 400ms post stimulus onset. In contrast, both the unpredicted and unpredicted-low-cloze waveforms showed more negative amplitudes which appeared to diverge from the predicted responses at approximately 200ms. At 300ms post-stimulus, the amplitude of these two negativities then appeared to diverge, with unpredicted targets showing a smaller and earlier negativity than the unpredicted low-cloze targets. Finally, following these negative components, both unpredicted and low-cloze targets showed a late positive shift which appeared to be maximal over frontal and left hemisphere electrode sites.

To assess these differences statistically, we performed a set of repeated-measures analyses of variance (ANOVA). To analyze the topographic distribution of these effects over the scalp we performed two separate ANOVAs for each comparison, one at Midline electrode sites (AFz, Fz, Cz, Pz, POz) which included a 5-level factor of Anteriority, and one at Lateral electrode sites which included a three-level factor of Anteriority: Frontal (FP1/2, F7/8, F3/4), Central (FC5/6, FC1/2, C3/4, CP1/2, CP5/6) and Posterior (T5/6, P3/4, O1/2), and a two-level factor of Hemisphere (Left, Right). For each analysis, any significant effects of the three-level Condition factor (predicted, unpredicted, low-cloze) were followed up by pair-wise comparisons. In some cases, multiple regression analyses were also used to determine which item-level factors could uniquely explain the observed ERP differences. We focus first on the negative components occurring in the N250 (200–300ms) and N400 (300–500ms) time windows. Then we turn to differences observed on the late frontal positivity (600–900ms).

3.2.1. Early time window (200–300ms)—In this analysis we were interested in determining the onset of the effects of prediction accuracy and contextual facilitation, and whether both effects could influence early lexical processing stages.[3] Mean amplitude values in this early, 200–300ms time window revealed a main effect of Condition (Midline: $F(2,46) = 24.08$, $p < 0.001$, Lateral: $F(2, 46) = 12.86$, $p < 0.001$), which was examined further using pairwise comparisons. Mean amplitudes in the predicted condition (3.98 μV) differed significantly from both the unpredicted condition (2.73 μV) and the low-cloze condition (2.61 μV), (Midline: $F(1,23) > 26$, all *ps* < 0.001, Lateral: $F(1,23) > 14$, all ps < 0.01). The size of this prediction effect did not differ significantly along the anterior-posterior axis, but was slightly more pronounced over the right hemisphere (Condition×Hemisphere interaction, Lateral: $F(1,23) > 8.0$, $p < 0.01$). Critically, in this early time window, no differences were observed between the unpredicted and low-cloze waveforms, Midline: $F(1,23) < 1$, Lateral: $F(1,23) < 1$.

3.2.2. N400 window (300–500ms)—We also analyzed mean amplitude values between 300 and 500ms where the N400 effect is typically maximal. This analysis revealed a highly significant effect of Condition (Midline: $F(2,46) = 136.15$, $p < 0.001$, Lateral: $F(2, 46) = 134.70$, $p < 0.001$). Amplitudes in the predicted condition (6.81 μV) differed significantly from both the unpredicted (3.44 μV) and the low-cloze condition (0.73 μV), (Midline: $F(1,23) > 90$, all *ps* < 0.001, Lateral: $F(1,23) > 100$, all *p*s < 0.001). In this later time window, the difference between unpredicted and low-cloze words was also highly significant (Midline: $F(1,23) = 87.91$, $p < 0.001$, Lateral: $F(1, 23) = 80.86$, $p < 0.001$). For each of these comparisons, we observed a significant Condition × Anteriority × Hemisphere interaction in the Lateral site analysis, $F(2,46) > 3.9$, all *p*s < 0.05, indicating that these differences were largest over central-parietal and right hemisphere electrode sites.

3.2.3 Latency analyses of difference waveforms—The differential pattern of results observed for the 200–300ms and 300–500ms time-windows suggests that successful lexical prediction may exert a relatively early influence on word processing, while other sentence-level factors may be somewhat delayed in comparison. To gain a better understanding of the time-course of these effects, two separate difference waveforms were calculated. To isolate the effects of prediction, we performed a point-by-point subtraction of the predicted medium-cloze waveforms and unpredicted medium-cloze waveforms. Note that, because these two conditions were time-locked to the same critical words embedded in the same sentence contexts, any activity in the resulting difference waveform is likely due to the effect of *prediction* accuracy. In addition, we isolated the effects of sentence context (independent of prediction) by subtracting unpredicted medium-cloze waveforms from unpredicted low-cloze waveforms. For simplicity, we will refer to this difference as a *contextual facilitation*, or *context* effect.

These difference waves revealed two distinct central-parietal negativities (see Figure 3). While the peak amplitudes of these effects were comparable[4], the two waveforms differed substantially in their overall time course. For each subject, we calculated the peak and onset

[3]Similar analyses performed in earlier 0–100ms and 100–200ms time windows revealed no significant Condition main effects or interactions. There were also no significant ERP differences on the two words preceding the final critical word.

[4]We compared peak amplitudes for the prediction and context effects between 150 and 650ms. This analysis revealed somewhat larger peak amplitudes for the prediction effect at Lateral electrode sites (−6.97 μV vs −5.97 μV), $F(1,23) = 4.46$ $p < 0.05$. These differences appeared to be most pronounced over the left hemisphere, Condition × Hemisphere: $F(1,23) = 9.64$, $p < 0.01$.

latency of these difference waves using a central-parietal cluster of electrodes (PZ, POZ, CP1, CP2, P3, P4), within a time range of 150–650ms. On average, the peak of the prediction effect occurred at 380ms (SD = 38), while for the context effect this point occurred at 480ms (SD=33). This timing difference (100 ± 15 ms) was highly significant, $F(1,23) = 191.17$, $p < 0.001$, $d = 5.77$, and was apparent in the difference waveforms from each of the 24 subjects. Onset latencies for these difference waves were determined by calculating a 20% peak latency measure for each subject over the same time window and electrode cluster. On average, the prediction effect reached this onset criterion at 267ms, SD = 34, while for the context effect the onset occurred at 352ms, SD=42. This difference (86 ± 20ms) was also significant, $F(1,23) = 81.30$, $p < 0.001$, $d = 3.76$.

3.2.4. Topographic analyses of difference waveforms—To compare the overall scalp distributions of the prediction and contextual facilitation effects, we calculated mean amplitude within a 100ms window, centered on the peak of each difference wave (prediction: 330–430ms, context: 430–530ms). A vector scaling procedure was used to normalize the size of these effects across subjects (McCarthy & Wood, 1985), and these values were entered into separate Midline and Lateral ANOVAs. At Midline sites we observed a main effect of Anteriority, $F(4,92) = 59.30$, $p < 0.001$, but no Anteriority×Condition interaction, $F(4,92) = 1.01$, $p = 0.35$, which suggests a common central-parietal maximum for both effects. In contrast, the Lateral analysis produced a three-way interaction between Condition, Hemisphere, and Anteriority, $F(2,46) = 7.84$, $p = 0.003$. This interaction was driven by a significant, right-hemisphere maximum for the context effect (−4.39 μV left, −5.44 μV right, $F(1,23) = 30.67$, $p < 0.001$), and a more bilateral distribution for the effect of prediction accuracy (−5.92 μV left, −5.93 μV right, $F < 1$).

In addition to these subtle topographic differences between the prediction and context effects, we also examined potential differences in the distribution of the prediction effect across the N250 and N400 time windows (see Figure 4). The motivation for this analysis was to determine whether activity in these two epochs might be generated by partially non-overlapping neural sources. We performed a vector-scaled analysis on the amplitude of the prediction effect, including our standard topographic factors and a two-level factor of Time (200–300ms vs 300–500ms). This resulted in a significant Time × Anteriority interaction (Midline: $F(4,92) = 17.82$, $p < 0.001$, Lateral: $F(2,46) = 4.39$, $p < 0.05$). In the early time window the prediction effect was largest at medial sites, with similar effects occurring at frontal (−1.38μV) and posterior (−1.65μV) electrodes, $F < 1$. In the later time window, the prediction effect instead showed a reliable posterior focus (−0.08μV frontal vs. −6.95μV posterior), $F(1,23) = 179.42$, $p < 0.001$.

In sum, the results show that incorrect predictions produced an early negativity that began approximately 200ms post-target onset. Initially, this component displayed a widespread scalp distribution largest at medial electrodes, which eventually resolved into a central parietal negativity reminiscent of the N400. When analyzing the effects of sentence context independent of prediction, low-cloze targets also produced a large, central-parietal negativity, potentially reflecting differences in contextual facilitation or integration difficulty. This effect was more pronounced over the right-hemisphere, and its onset and peak latencies were delayed by approximately 100ms relative to the effects of prediction.

3.2.5. Source of contextual facilitation effects: regression analyses—While we believe the early *prediction* related negativity can safely be attributed to the effects of prediction accuracy, the cause of the late parietal negativity, or *contextual facilitation* effect, is less clear-cut. As originally defined, the primary difference between our low-cloze and medium-cloze passages was the degree to which they constrained the final, critical word. Although this comparison only included *unpredicted* lexical items, substantial ERP differences still remained which distinguished target words appearing in these two contexts. These differences must have been caused by some feature(s) present in the preceding discourse, but several distinct processes may have contributed. We used a multiple regression approach to help distinguish between several possible candidates, which are outlined below.

As discussed previously, the low and medium-cloze passages differed in overall *plausibility*, and therefore this later negativity may simply reflect increased integration demands occurring for the low-cloze targets. In addition, although care was taken to match the lexical content immediately preceding the critical item, *lexical associates* present earlier in the discourse may have also contributed to these late ERP differences. Finally, incorrect predictions generated during the medium-cloze passages may have shared a larger number of semantic features with the actual target words than predictions generated during low-cloze passages. For example, the anticipated word ‘book’ may have facilitated processing of the actual target ‘novel’ by virtue of its many shared semantic features (“has pages”, “can be read”), while the anticipated word ‘dish’ would not (Roland, Yun, Koenig, & Mauner, 2012). This difference in overall *semantic overlap* may have also played an important role.

To determine if any of these mechanisms (*plausibility*, *lexical association*, and *semantic overlap*) could predict the amplitude of the observed context effect, we first quantified each of these variables for each experimental passage. To assess passage *plausibility*, we used the offline plausibility ratings generated from two groups of 36 UC Davis undergraduates (see section 2.2). To calculate a measure of *lexical association*, we used Latent Semantic Analysis (LSA) values obtained from the “one-to-many function” available at lsa.colorado.edu. The LSA cosine between each critical word and the entirety of its preceding passage was calculated using the General Reading – Up to First Year of College topic space with 300 factors (Landuaer & Dumais, 1997). To calculate a measure of *semantic overlap*, we first used the results of our cloze norming procedure to determine the most likely completion of each low-cloze passage and the next-best completion of each medium-cloze passage (for examples, see Appendix). We then used LSA to assess the degree of semantic overlap between each alternate completion and the actual final word (e.g. book-novel = 0.30, dish-novel = 0.04). As expected, target words in the medium-cloze passages showed a higher degree of *lexical association* with words in the preceding context (medium cloze: mean = 0.36 SD = 0.13, low cloze: mean = 0.30, SD = 0.13, $t(179) = 9.48$, $p < 0.001$). Medium-cloze passages also showed a higher degree of *semantic overlap* (medium-cloze: mean = 0.45 SD = 0.22, low-cloze: mean = 0.19, SD = 0.15, $t(179) = 14.19$, $p < 0.001$).

To assess the unique contribution of these three factors on the late ERP negativity, we averaged across subjects to create individual by-items ERPs for each passage, which were time-locked to the final word. Because we wished to assess the influence of contextual factors in the absence of lexical pre-activation, this analysis included only incorrectly predicted trials, and averaging was performed separately for medium-cloze and low-cloze targets. Following artifact rejection, there were 7.8 trials on average for each of the 360 items. As in our previous analyses, mean amplitude values were calculated between 430–530ms for a cluster of central-parietal electrodes (Pz, POz, CP1, CP2, P3, P4). Finally, a linear multiple regression analysis was performed, predicting the amplitude of the late negativity across passages.[5]

The model revealed a significant influence of *plausibility*, $b = -1.24$, $t = -6.55$, $p < 0.001$, with larger negative amplitudes for implausible passages. There was also an independent effect of *semantic overlap*, $b = 6.41$, $t = 4.86$, $p < 0.001$, with smaller negative amplitudes for targets with a related alternate completion. The effect of *lexical association* did not approach significance, $t = -0.61$, $p = 0.55$. Overall, the model was significant, $F(3,356) = 39.55$, $p < 0.001$, with an adjusted $R^2 = 0.25$ (see Table 1). Scatter plots, representing these two significant effects are displayed in Figure 5. Overall, this analysis suggests that, for unpredicted lexical items, both coherence with the preceding discourse and activation of overlapping semantic features reduced the amplitude of this late, central-parietal negativity. In contrast, low-level lexical association did not play a unique role in modulating this component.

3.2.6. Late Frontal Positivity—In addition to the negative components present at parietal sites, visual inspection of the waveforms in Figure 2 revealed systematic differences in the amplitude of a late frontal positivity. This component was of particular *a priori* interest, because it has previously been linked to manipulations of sentence constraint and has been hypothesized as a marker for the “costs” of incorrect prediction. For the PNP we again performed ANOVAs over Midline and Lateral electrode sites using mean amplitude values between 600–900ms in each condition (predicted, unpredicted, low-cloze). This analysis revealed a main effect of Condition (Midline: $F(2,46) = 6.45$, $p = 0.007$, Lateral: $F(2, 46) = 6.02$, $p = 0.011$), as well as significant interactions with Anteriority (Midline: $F(8,184) = 8.91$, $p < 0.001$, Lateral: $F(4, 92) = 3.16$, $p = 0.036$) and Hemisphere (Lateral: $F(2,46) = 20.11$, $p < 0.001$), indicating that differences in the PNP were maximal at frontal and left hemisphere electrode sites (see Figure 4). For the follow-up pair-wise comparisons, we confined our analyses to a frontal cluster of 6 electrode sites where the PNP was maximal (AFZ, FZ, FP1, FP2, F3, F4). This analysis revealed that unpredicted medium-cloze (2.73 μV) and unpredicted low-cloze words (3.66 μV) both showed more positive amplitudes than predicted final words (1.40 μV): $F(1,23) = 7.19$, $p = 0.013$ and $F(1,23) = 14.49$, $p < 0.001$, respectively. The difference between unpredicted low-cloze and unpredicted medium-cloze targets was also significant: Condition $F(1,23) = 4.71$, $p = 0.04$, Condition × Electrode $F(5,115) = 6.73$, $p = 0.001$.

[5]These three predictors were weakly to moderately correlated: plausibility and lexical association, $r = -0.13$, plausibility and semantic overlap $r = -0.44$, lexical association and semantic overlap $r = 0.34$. Variance Inflation Factors for our multiple regression analyses never exceeded 1.38.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

While these analyses suggest that both prediction accuracy and contextual facilitation contribute to the amplitude of the PNP, we wished to confirm that these effects were modulating the same underlying component. To compare the topographic distribution of these effects, we analyzed mean amplitude values between 600–900ms for the two difference waves. As in our other topographic analyses, these values were vector-scaled and analyzed separately for Midline and Lateral electrode sites. This analysis revealed no significant interactions between Condition and any of our topographic factors, suggesting these two PNP effects were likely generated by overlapping neural sources (see Figure 4).

In sum, unpredicted sentence completions produced a larger PNP over frontal and left hemisphere electrode sites. This finding supports the hypothesis that the PNP reflects the costs associated with detecting or resolving an incorrect prediction. In addition, this component was larger, not just for unpredicted targets, but also for words appearing in less supportive sentence contexts. To determine which aspects of the preceding context may have been driving these differences, we again used multiple regression techniques to predict the amplitude of the PNP on an item-by-item basis as a function of *plausibility*, *lexical association*, and *semantic overlap*. We first calculated mean amplitudes between 600 and 900ms after the presentation of the final critical word, using the same cluster of frontal electrode sites. We again used only unpredicted trials, averaging across subjects for each of the 360 passages. This analysis revealed a significant effect of *plausibility*, $b = 0.52$ $t = 3.61$, $p < 0.001$, with larger positive amplitudes occurring for more implausible passages. In contrast, *semantic overlap*, $t = 0.59$, and *lexical association*, $t = -0.81$ had no significant effect on the amplitude of the PNP (Table 2).

4. General Discussion

This study sought to disentangle the influence of lexical prediction from other forms of contextual facilitation during reading. Using participant self-reports we compared electrophysiological responses to predicted and unpredicted target words appearing in supportive and non-supportive discourse passages. This approach revealed temporally distinct effects of prediction accuracy and discourse context on lexical processing. Prediction showed the earliest influence on ERP amplitudes, with predicted targets showing a reduction in the N250. In contrast, the effects of contextual facilitation were delayed by approximately 100ms, and persisted for longer into the ERP epoch. In addition, at frontal electrode sites, both prediction accuracy and discourse context influenced a late positive component, the PNP, which has previously been linked to the costs of disconfirmed predictions. We will first discuss the relevance of these two negative components and later turn to the functional interpretation of the frontal positivity.

4.1. Lexical Prediction Effects

In this study, prediction had a rapid influence on lexical processing, modulating ERPs approximately 200ms after the appearance of the critical, sentence-final word. The early portion of this effect, occurring between 200 and 300ms did not have the characteristic distribution of the N400; instead it showed a broad scalp distribution that was largest at medial sites, with roughly equal amplitudes along the anterior-posterior axis. In contrast, later portions of the prediction-related negativity showed a pronounced central-parietal distribution, which is consistent with a reduction in the N400 component following a successful lexical prediction.

The rapid onset observed for this prediction effect suggests that pre-activation of specific word candidates can influence early stages of lexical processing. In the auditory sentence processing domain, an effect with a similar onset latency and scalp distribution has also been shown in response to unpredictable target words whose initial phonemes differ from an anticipated continuation (Connolly & Phillips, 1994, Van der Brink, Brown & Hagoort, 2001). This N200/N250 effect has been described as indexing the detection of a phonological mismatch, or alternatively the initiation of lexical access for a new, unanticipated target word. In the visual modality, studies employing the masked priming technique (e.g. Holcomb & Grainger, 2006), have shown that repetitions of an upcoming word can also produce ERP effects with a similar polarity and onset latency. The visual N250 effect has been hypothesized to reflect facilitation of early orthographic and phonological processing and also shows a more widespread scalp distribution, distinct from the N400. Considering these similarities, we suggest that the early prediction-related negativity in the current study may reflect a functionally similar mechanism to the auditory N200/N250 as well as the visual N250 effect. These three effects may share a distinct polarity, time-course, and scalp distribution through facilitation of a common lexical processing stage (for a similar suggestion regarding early stages of the N400 see Kuperberg, in press).[6] Previous studies have also observed a dissociation between early and late anticipatory effects during sentence processing. In a study by DeLong and colleagues (2005), unpredicted sentence continuations that differed in form but not meaning (“a” vs “an”) produced an early negativity in the N250 time range (approximately 250–350ms) but showed no differences during later stages of the N400 (350–500ms). This finding, which is consistent with the latency differences observed in the current study, provides additional support for a distinction between form-based and meaning-based pre-activation.

While a growing collection of studies has observed lexical prediction effects in the 200–300ms time-range, other ERP and MEG studies have also observed effects of lexical prediction on even earlier visual components such as the P1 (Kim & Lai, 2012), or its magnetic equivalent the M100 (Dikker & Pylkkänen, 2011). While standard, cloze-probability manipulations do not appear to modulate these early visual components in ERPs, different task manipulations or the increased sensitivity of visual MEG recordings may be critical for observing these effects. Clearly, an important question for future research will be determining how and when lexical prediction influences visual processing.

In later time windows, we observed large processing differences (approximately 6 μV) between predicted and unpredicted target words, which suggest that subjects experienced

[6]Other investigators have suggested that these early portions of the N400 effect are driven by modulations of the P2 component (e.g. Federmeier, Mai & Kutas, 2005; Federmeier et al., 2007). Unlike attentional P2 effects, which typically occur at anterior sites (e.g. Luck & Hillyard, 1994, figure 2), our early prediction effect showed a more widespread scalp distribution. In addition, while spatial attention and target saliency typically *increase* evoked visual responses, target predictability has been shown to *reduce* activity in early visual areas (Alink, et al., 2010). We believe that this portion of the prediction effect (200–300ms) is most consistent with a reduction in the N250 (a component thought to reflect early word form processing).

robust facilitation following an accurate prediction. While the presentation of a meaningful stimulus typically results in some observable N400 activity, a close inspection of the predicted, medium-cloze waveforms (Figure 2) reveals a complete absence of a typical N400 peak within the 300–500ms time window. In previous studies, the absence of an N400 peak has also been observed in response to very high-cloze, sentence-final words (cloze = 0.88, Van Petten & Kutas, 1991), as well as highly predictable targets in prime - target, antonym pairs (e.g. black - white, Kutas & Iragui, 1998). In addition to attenuating the N400, highly predictive contexts have been shown, both behaviorally and electrophysiologically, to eliminate the effects of many variables which typically influence lexical processing, including word frequency (Van Petten & Kutas, 1990), concreteness (Schwanenflugel & Shoben, 1983, Holcomb, et al., 1999), and orthographic neighborhood (Molinaro, et al., 2010). Apparently, following a correct lexical prediction, orthographic analysis and lexical selection processes are either complete, or have been pre-activated to such a degree that these important lexical variables no longer produce a measurable response. In the current study, the large processing differences observed between predicted and unpredicted targets is a testament to the rapid and robust facilitation provided by accurate prediction. Moreover, these results suggest that there is no simple one-to-one mapping between surrounding semantic constraints and the benefits of lexical prediction. It appears that, even when the surrounding discourse is held constant, the benefits of prediction can vary substantially on a trial-by-trial basis, depending on which lexical item(s) were ultimately selected for pre-activation.

While we have interpreted the prediction effects observed in this study as reflecting facilitation at the orthographic (N250) and semantic (N400) level, it is important to note that these results may be additionally driven by feedback- or error-related processing. In non-linguistic paradigms, such as gambling tasks, participants show differential ERP signals in response to positive and negative feedback (Hajcak, Moser, Holroyd & Simons, 2006). Error-related processing of this kind has been tied to activity in anterior cingulate cortex (ACC) and is believed to play an important role in reinforcement learning (Holroyd & Coles, 2002). In this way, our unpredicted sentence continuations may have served as a form of negative feedback which in turn triggered more domain-general, error processing mechanisms. Some have recently theorized that predictive error signals of this kind may play a fundamental role in both sentence comprehension (Levy, 2008) and production (Jaeger & Snider, 2013). Currently, it is an open question to what extent these early, prediction-related ERP differences reflect lexical facilitation, error detection, or a combination of the two. Nonetheless, we can conclude from the onset latency of this component that the brain is *uniquely* sensitive to the effects of lexical prediction during the initial stages of word processing.

Another important question is whether the presence of an overt prediction task may have modulated ERP responses in the current experiment. While the prediction task did place additional demands on subjects, we nonetheless observed substantial effects of both semantic overlap and contextual plausibility on the N400. This suggests that participants were actively reading for comprehension and were successfully integrating sentence-final words into the preceding discourse context. While an overt task likely amplifies the effects of anticipatory processing during reading, we believe the mechanisms used to generate lexical predictions in the current study are similar to those occurring under more naturalistic conditions (e.g. Van Berkum, et al., 2005). Overall, we believe this prediction task can provide a complementary method for assessing the time-course of anticipatory processing during reading.

4.2. Contextual Facilitation Effects

Approximately 100ms after the onset of the prediction-related negativity, the effects of contextual facilitation also began to influence the ERP record. From the amplitude reductions observed on the N400 component, it appears that, even after controlling for the effects of prediction accuracy, contextual support still plays an important role in facilitating lexical processing. Our multiple regression analyses revealed that N400 amplitude in this later time range was strongly correlated with two independent factors: 1) the degree of shared semantic overlap between the presented target and the next-best completion of the passage, and 2) the rated plausibility of the passage as a whole.

In prior research, the link between semantic overlap and the N400 has been well established. In these studies, participants read constraining sentences with final words that were unpredictable, yet semantically (Kutas & Hillyard, 1984) or categorically (Federmeier & Kutas, 1999) related to an alternative “best completion”.[7] As in the current study, semantically related targets, even those which were anomalous or unpredictable in context, showed smaller N400 amplitudes. This suggests that the N400, to some extent, indexes the effort of accessing the semantic features of a lexical item, and that this process can become primed through the pre-activation of related target words. Critics of anticipatory processing in language contexts have often claimed that prediction mechanisms would be prohibitively costly, especially considering how often specific predictions would be incorrect (Jackendoff, 2002, Morris, 2006). It may be the case though, that by committing to a lexical item (e.g. *“book”*) and activating its network of related concepts (pages, reading, editors), a reader may improve her comprehension of the upcoming passage, even if the word “book” itself never appears. The effects of anticipatory, semantic overlap may help to explain how the benefits of lexical anticipation can often outweigh the costs.

4.3. Time course of effects

One of the most striking results of this study was the strong temporal dissociation between the effects of prediction and contextual facilitation. This timing difference may suggest that there is no single point during lexical processing when all potential constraints affecting word processing simultaneously come to bear. Instead, lexical prediction may enjoy a unique *temporal primacy* in facilitating word access. Additional support for this hypothesis comes from Lau, Holcomb and Kuperberg (2013) who presented prime-target pairs while manipulating the overall proportion of semantically related stimuli within an experimental block. In the high-proportion block, when lexical prediction was strategically valid, N400 priming effects were larger and were accompanied by an earlier, widespread effect between

[7]Although latency analyses were not performed, the results of Kutas and Hillyard, 1984 (Figure 2) and Thornhill and Van Petten 2012 (Figure 4) appear to show a temporal delay between the effects of cloze probability and semantic overlap, consistent with the current results.

200–300ms. A similar result was obtained in a word priming study by Luka and Van Petten (2014). In this study, strongly related words produced an early facilitation effect in the 200–300ms time window when presented sequentially. Critically, these effects disappeared when word pairs were presented on the screen simultaneously, presumably because participants had no time to generate a lexical prediction.

How can we interpret these results in the light of evidence suggesting that higher-level linguistic processes (message-level plausibility, pragmatic information, world knowledge) can have a rapid influence on ERPs, often within 200ms post-target (e.g. Hagoort & van Berkum, 2007)? Considering that rich sources of contextual information, such as speaker identity, are available long before the appearance of a particular target word, this information could potentially influence the selection and pre-activation of specific lexical candidates as a discourse unfolds. Indeed, prediction can provide a plausible mechanism for explaining these early ERP effects. While it is unlikely that a listener could calculate the pragmatic appropriateness of a young child saying “Every night I drink some *wine*” within 200ms post target-onset (before the utterance itself is even complete), we speculate that the detection of a phonological mismatch (e.g. between the target word “wine” and a predicted word “juice”) could more plausibly occur on this time scale. In this way, the pre-activation of specific lexical items would produce a common neural signature across a wide variety of semantic and pragmatic constraints.

4.4. Post-N400 Positivity (PNP)

Beyond the N400 window, the difference waves in Figure 3 reveal a biphasic response, with unpredicted items showing a larger positivity that was most pronounced at frontal and left-hemisphere electrode sites. This component was similar in both its time-course and topography to frontal positivities elicited in previous studies with standard cloze-probability manipulations (Federmeier et al., 2007, DeLong, Urbach, Groppe & Kutas, 2011, Thornhill & Van Petten, 2012). The finding that this component is directly modulated by prediction accuracy provides strong evidence for a link between the PNP and the resolution of disconfirmed predictions. If this frontal PNP simply reflects either 1) the detection of a disconfirmed prediction or 2) the cost of inhibiting incorrectly selected lexical items, PNP amplitude should vary only as a function of prediction accuracy, with no additional influence of the preceding sentence context. Contrary to this hypothesis, we observed a graded response across our three experimental conditions, with both prediction accuracy and contextual support modulating the PNP.[8]

One potential explanation for these results is that, in constraining discourse contexts, predicted lexical items are both pre-activated and *integrated* into the preceding context as soon as they become available. Under these circumstances, encountering an incorrect prediction would require both lexical processing of the new target and a partial *revision* of the current discourse model. This context-updating hypothesis is consistent with our finding

[8]In an ERP study, Lau, Holcomb and Kuperberg (2013) employed a relatedness proportion manipulation to modulate participants' prediction strategies. This manipulation caused an early latency shift in the N400, similar to the prediction effects observed in the current study. In contrast, the authors found *no* evidence of a frontal positivity triggered by unpredicted target words. The absence of a frontal positivity in single word contexts suggests that the PNP may reflect brain mechanisms specific to sentence-level processing.

that PNP amplitude was significantly correlated with the overall plausibility of unpredicted target words. This suggests that items with less plausible completions (e.g. a chef working on his recent novel) required more elaborate revisions or inferences to successfully integrate the sentence-final word, and this difficulty was then reflected in the amplitude of the PNP. In contrast, this late positive component was unaffected by the degree of semantic overlap, a factor known to influence both lexical processing difficulty and N400 amplitude (see also Thornhill & Van Petten, 2012). This dissociation further supports the idea that the PNP reflects a distinct post-lexical, discourse revision mechanism.

4.5 Mechanisms underlying the Prediction Effect

To date, the evidence in support of lexical prediction has not determined whether lexical candidates are activated probabilistically (with activation weights equal to their cloze probabilities) or if these candidates are selected, one-at-a-time. Probabilistic accounts are most consistent with connectionist and constraint based approaches in the psycholinguistic literature, but recently, support for single activation accounts has also emerged, motivated by the functional interconnections between our language production and comprehension systems (Pickering & Garrod, 2007). As incoming language is processed, the language production system could be used to select and activate a viable continuation of the current discourse. After selecting this continuation, processing costs in response to the actual stimulus would be minimized, resulting in a form of implicit repetition priming.

Currently, evidence for resolving these conflicting accounts is scant. The linear correlation between N400 amplitude and cloze probability has been previously offered as supporting evidence for probabilistic activation (Delong, Urbach & Kutas, 2005), but this same experimental result can be produced under a single-activation accounts well, simply as a byproduct of multi-subject averaging. For example, in an averaged waveform, it is ambiguous whether a critical word with a 30% cloze probability was pre-activated to a strength of 30% by all 10 participants or to 100% by only 3 participants; the same degree of facilitation would ultimately be observed under both scenarios.

In the current study, if activity in the earliest post-target time window (200–300ms) was reflective of early orthographic and lexical processing, then probabilistic activation accounts should predict facilitation during this period as a function of the preceding context. According to this account, in the supportive “author” context, the lexical items ‘book’ and ‘novel’ should both become pre-activated at the orthographic level, presumably outside conscious awareness. The fact that we only observed effects of binary prediction accuracy (predicted vs unpredicted) in this early time window, and saw no separate influence of discourse context, suggests two possibilities. Either, 1) probabilistic pre-activation of the non-selected target was suppressed at some point prior to the onset of the critical word, or 2) probabilistic pre-activation did not occur, and the act of selection itself was the only source of early, form-based priming. Ultimately, the precise mechanism underlying lexical prediction remains an open question for future research, but the present evidence *does* suggest that conscious anticipation of a single target word can produce robust and rapid facilitation during lexical processing.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Appendix

A representative sample of medium-cloze (M) and low-cloze (L) passages. Final critical words are italicized here for emphasis. Displayed in parenthesis are the "next best" (M) or "best" (L) completions for each sentence pair as determined by our cloze test. Note that these alternative completions were never presented as experimental targets and that each participant saw only one passage from each pair.

M: The author is writing another chapter about the fictional detective. To date, he thinks it will be his most popular *novel*. (book)

L: Everyone congratulated the chef on all his hard work. To date, he thinks it will be his most popular *novel*. (dish)

M: Grandma Tootsie was walking on the icy sidewalk. Her grandson supported her arm so she would not *slip*. (fall)

L: The Minister had just suffered the loss of her husband. Her grandson stood by her so she would not *slip*. (cry)

M: The old school teacher wanted to draw the diagram up on the board. Unfortunately, he could not find his *chalk*. (marker)

L: The lawyer's pencil broke, and he couldn't finish writing his idea. Unfortunately, he could not find his *chalk*. (pen)

M: John had been working all day and flopped down on the bed. He definitely needed to get some *rest*. (sleep)

L: Because of John's hectic schedule, his pantry was completely empty. He definitely needed to get some *rest*. (food)

M: The carpenter climbed up a ladder to repair the house. A recent wind storm had blown off the *roof*. (shingles)

L: The kids could not get any of their satellite TV channels. A recent wind storm had blown off the *roof*. (dish)

M: Dan walked past the playground during recess. He heard the laughter of all the cute little *children*. (kids)

L: Dan visited the petting zoo at the county fair. He heard the squeals of all the cute little *children*. (pigs)

M: Hilda was interested in politics. She hoped one day to become the mayor of her *city*. (town)

L: At school, Hilda's classmates were having an election. She hoped that she would become the mayor of her *city*. (class)

M: Tim wouldn't eat any kind of seafood with claws. He especially didn’t like *crab*. (lobster)

L: The chef hated working with raw, ground meat. He especially didn’t like *crab*. (beef)

M: The CEO had strange behavior and violent outbursts. This caused his coworkers to believe he was going *crazy*. (insane)

L: Dave finished his assignment and quickly left the building. All of his coworkers assumed that he was going *crazy*. (home)

M: Near the intersection, Bill was texting. He did not notice that the light had turned *red*. (green)

L: Bill was sick of the faulty wiring in his apartment. He did not notice that the light had turned *red*. (off)

M: David has to pick between the two alternatives before Monday. He doesn't think he'll be able to *decide*. (choose)

L: David needs to leap across the wide chasm if he wants to escape. He doesn't think he'll be able to *decide*. (jump)

M: To test them, Steve smelled the clothes in the laundry basket. He realized that they were all *dirty*. (clean)

L: Steve wondered why the room was so dark, so he checked the blinds. He realized that they were all *dirty*. (closed)

M: Debbie wanted a long crunchy vegetable to dip into the ranch dressing. She decided to buy some *celery*. (carrots)

L: At the store, Debbie needed a treat that would quiet her fussy children. She decided to buy some *celery*. (candy)

M: Jim thought he heard someone call his name. He looked back through the open *door*. (window)

L: The cross country runner loved exploring new places. He ran ahead through the open *door*. (field)

M: Debbie suddenly fainted in the crowded room. She couldn’t fall to either side so she ended up falling *forward*. (backward)

L: The puppy had a long day running in circles. She was so exhausted from playing that she ended up falling *forward*. (asleep)

M: Bill straightened his tie on the way to the church. He only wore dark suits when attending a *funeral*. (service)

L: The professor hated the hot weather outside. He only wore long pants when attending a *funeral*. (lecture)

M: The researcher built a time machine. He has always wanted to travel into the *future*. (past)

L: The hiker wanted to escape from all of his problems. He has always wanted to travel into the *future*. (mountains)

M: Since the accident, Bill can't even hold a pencil. He has a horrible pain running down his *arm*. (hand)

L: Since the adoption, Bill's cat has been constantly trying to escape. He can see it now running down his *arm*. (street)

M: The candlesticks they brought to the table matched the duchess' pendant. Both had emeralds and were made out of *gold*. (silver)

L: The archeologist brought two ancient vases from the prehistoric cave. Both had cracks and were made out of *gold*. (clay)

M: Alex loved playing string instruments, but was sick of the viola. He wanted his mother to buy him a *guitar*. (violin)

L: Alex had gotten bored of his old Play station. He wanted his mother to buy him a *guitar*. (xbox)

M: Thomas didn't like the temperature of his drink. He thought it was much too *hot*. (cold)

L: Thomas didn't like the look of the water. He thought it was much too *hot*. (dirty)

M: It was surprising to hear the African locals talking about their royal monarchy. Tim did not know that their country had a *king*. (queen).

L: The leader explained to the visiting delegates why all their crops were dried out. Tim did not know that their country had a *king*. (drought)

M: When Jason reached the intersection he flicked on his turn signal. At the last minute, he decided to turn *left*. (right)

L: Jason doesn't think that he should go any farther into the cavern. After lots of hesitation, he decided to turn *left*. (around)

M: At camp, Ken was surprised by the amount of food on his plate. He would have served himself a lot *less*. (more)

L: At the restaurant Ken hated how slow all of the waiters were. He would have served himself a lot less. (faster)

M: At the party, Denise changed the volume on the stereo. Before it was obviously too *loud*. (quiet)

L: At the party, Denise changed the settings on her hot tub. Before it was obviously too *loud*. (hot)

M: People waste so much money during birthdays and holidays. Amanda hates when people buy her too many *gifts*. (presents)

L: Now there are four dogs in the house and three parakeets. Amanda hates when people buy her too many *gifts*. (pets)

M: Drew struggled to open the door the wrong way. After reading the sign, he realized that he needed to *push*. (pull)

L: Drew was running late, but the train was completely full. After reading the sign, he realized that he needed to push. (wait).

M: When Peter checked his account balance, he noticed the bank had made an error. He realized he was suddenly much *richer*. (poorer)

L: Peter's memory training had really been paying off. He realized he was suddenly much *richer*. (smarter)

M: Bill always had political ambitions at the national level. He hopes to become a member of the United States *Congress*. (Senate)

L: Bill is an excellent shot and wants to protect his country. He hopes to become a member of the United States *Congress*. (Army)

References

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. Stimulus predictability reduces responses in primary visual cortex. The Journal of Neuroscience. 2010; 30(8):2960–2966. [PubMed: 20181593]

Altmann G, Mirković J. Incrementality and prediction in human sentence processing. Cognitive Science. 2009; 33(4):583–609. [PubMed: 20396405]

Bransford JD, Johnson MK. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior. 1972; 11(6):717–726.

Connolly JF, Phillips NA. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. Journal of Cognitive Neuroscience. 1994; 6(3):256–266. [PubMed: 23964975]

Delong KA, Urbach TP, Groppe DM, Kutas M. Overlapping dual ERP responses to low cloze probability sentence continuations. Psychophysiology. 2011; 48(9):1203–1207. [PubMed: 21457275]

DeLong KA, Urbach TP, Kutas M. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nature neuroscience. 2005; 8(8):1117–1121.

DeLong KA, Quante L, Kutas M. Predictability, plausibility and two late ERP positivities during written sentence comprehension. Neuropsychologia. 2014 In Press.

Elman JL. Finding Structure in Time. Cognitive Science. 1990; 14(2):179–211.

Federmeier KD, Kutas M. A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. Journal of Memory and Language. 1999; 41(4):469–495.

Federmeier KD, Mai H, Kutas M. Both sides get the point: Hemispheric sensitivities to sentential constraint. Memory & Cognition. 2005; 33(5):871–886. [PubMed: 16383175]

Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. Multiple effects of sentential constraint on word processing. Brain Research. 2007; 1146:75–84. [PubMed: 16901469]

Foss DJ. A discourse on semantic priming. Cognitive Psychology. 1982; 14(4):590–607. [PubMed: 7140212]

Foss DJ, Ross JR. Great expectations: Context effects during sentence processing. The process of language understanding. 1983:169–191.

Gibson E, Bergen L, Piantadosi S. Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. PNAS. 2013; 110:8051–8056. [PubMed: 23637344]

Hagoort P, van Berkum J. Beyond the sentence given. Philosophical Transactions of the Royal Society B: Biological Sciences. 2007; 362(1481):801–811.

Hale, J. Proceedings of NAACL-2001. Stroudsburg, PA: Association for Computational Linguistics; 2001. A probabilistic early parser as a psycholinguistic model; p. 159-166.

Holcomb PJ, Grainger J. On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. Journal of Cognitive Neuroscience. 2006; 18(10): 1631–1643. [PubMed: 17014368]

Holcomb PJ, Kounios J, Anderson JE, West WC. Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1999; 25(3):721.

Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychological Review. 2002; 109(4):679. [PubMed: 12374324]

Jackendoff, R. Foundations of Language. New York: Oxford University Press; 2002.

Jaeger TF, Snider NE. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. Cognition. 2013; 127(1):57–83. [PubMed: 23354056]

Jordan TR, Thomas SM. In search of perceptual influences of sentence context on word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28(1):34–45.

Kamide Y, Altmann G, Haywood SL. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. Journal of Memory and Language. 2003; 49(1):133–156.

Kuperberg, GR. The Proactive Comprehender: What Event-Related Potentials tell us about the dynamics of comprehension. In: Miller, B.; Cutting, L.; McCardle, P., editors. Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension. Baltimore: Paul Brookes Publishing; (In Press).

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307(5947):161–163. [PubMed: 6690995]

Kutas M, Iragui V. The N400 in a semantic categorization task across 6 decades. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section. 1998; 108(5): 456–471.

Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. Trends in cognitive sciences. 2000; 4(12):463–470. [PubMed: 11115760]

Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological review. 1997; 104(2):211.

Lau EF, Holcomb PJ, Kuperberg GR. Dissociating N400 effects of prediction from association in single-word contexts. Journal of cognitive neuroscience. 2013; 25(3):484–502. [PubMed: 23163410]

Levy R. Expectation-based syntactic comprehension. Cognition. 2008; 106(3):1126–1177. [PubMed: 17662975]

Luck SJ, Hillyard SA. Electrophysiological correlates of feature analysis during visual search. Psychophysiology. 1994; 31(3):291–308. [PubMed: 8008793]

McCarthy G, Wood CC. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section. 1985; 62(3):203–208.

Miller GA, Heise GA, Lichten W. The intelligibility of speech as a function of the context of the test materials. Journal of experimental Psychology. 1951; 41(5):329. [PubMed: 14861384]

Molinaro N, Conrad M, Barber HA, Carreiras M. On the functional nature of the N400: Contrasting effects related to visual word recognition and contextual semantic integration. Cognitive Neuroscience. 2010; 1(1):1–7. [PubMed: 24168241]

Morris, RK. Handbook of psycholinguistics. Amsterdam; Boston: Elsevier, Academic Press; 2006. Lexical Processing and Sentence Context Effects.

Paller KA, Kutas M, Mayes AR. Neural correlates of encoding in an incidental learning paradigm. Electroencephalography and Clinical Neurophysiology. 1987; 67(4):360–371. [PubMed: 2441971]

Pickering MJ, Garrod S. Do people use language production to make predictions during comprehension? Trends in Cognitive Sciences. 2007; 11(3):105–110. [PubMed: 17254833]

Rayner K, Slattery TJ, Drieghe D, Liversedge SP. Eye movements and word skipping during reading: effects of word length and predictability. Journal of Experimental Psychology: Human Perception and Performance. 2011; 37(2):514. [PubMed: 21463086]

Roland D, Yun H, Koenig J-P, Mauner G. Semantic similarity, predictability, and models of sentence processing. Cognition. 2012; 122(3):267–279. [PubMed: 22197059]

Sanquist TF, Rohrbaugh JW, Syndulko K, Lindsley DB. Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. Psychophysiology. 1980; 17(6):568–576. [PubMed: 7443924]

Schwanenflugel PJ, Shoben EJ. Differential context effects in the comprehension of abstract and concrete verbal materials. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1983; 9(1):82–102.

Schwanenflugel PJ, Shoben EJ. The influence of sentence constraint on the scope of facilitation for upcoming words. Journal of Memory and Language. 1985; 24(2):232–252.

Schank, RC. Conceptual information processing. Amsterdam; New York: North-Holland, American Elsevier; 1975.

Swaab, TY.; Ledoux, K.; Camblin, CC.; Boudewyn, M. Language-related ERP components. Luck, Steven J.; Kappenman, Emily, editors. Oxford Handbook of Event-Related Potential Components; 2012.

Szewczyk JM, Schriefers H. Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. Journal of Memory and Language. 2013; 68(4):297–314.

Thornhill DE, Van Petten C. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. International Journal of Psychophysiology. 2012; 83(3): 382–392. [PubMed: 22226800]

Traxler MJ, Foss DJ. Effects of sentence constraint on priming in natural language comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2000; 26(5):1266.

van Berkum JJA, Hagoort P, Brown CM. Semantic Integration in Sentences and Discourse: Evidence from the N400. Journal of Cognitive Neuroscience. 1999; 11(6):657–671. [PubMed: 10601747]

van Berkum JJ, Brown CM, Zwitserlood P, Kooijman V, Hagoort P. Anticipating upcoming words in discourse: evidence from ERPs and reading times. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31(3):443.

Van Den Brink D, Brown CM, Hagoort P. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. Journal of Cognitive Neuroscience. 2001; 13(7):967–985. [PubMed: 11595099]

Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. Memory & Cognition. 1990; 18(4):380–393. [PubMed: 2381317]

Van Petten C, Kutas M. Influences of semantic and syntactic context on open-and closed-class words. Memory & Cognition. 1991; 19(1):95–112. [PubMed: 2017035]

Van Petten C, Luka BJ. Prediction during language comprehension: benefits, costs, and ERP components. International Journal of Psychophysiology. 2012; 83(2):176–190. [PubMed: 22019481]

Wicha NYY, Moreno EM, Kutas M. Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. Journal of Cognitive Neuroscience. 2004; 16(7):1272–1288. [PubMed: 15453979]

Highlights

- Participants read medium (50%) and low-cloze (>1%) words in discourse contexts.
- Neural activity was analyzed separately for predicted and unpredicted targets.
- Accurate predictions resulted in early differences on the N250 and N400.
- Effects of contextual support were delayed by approximately 100ms.
- Prediction and discourse plausibility both influenced a late, frontal positivity.

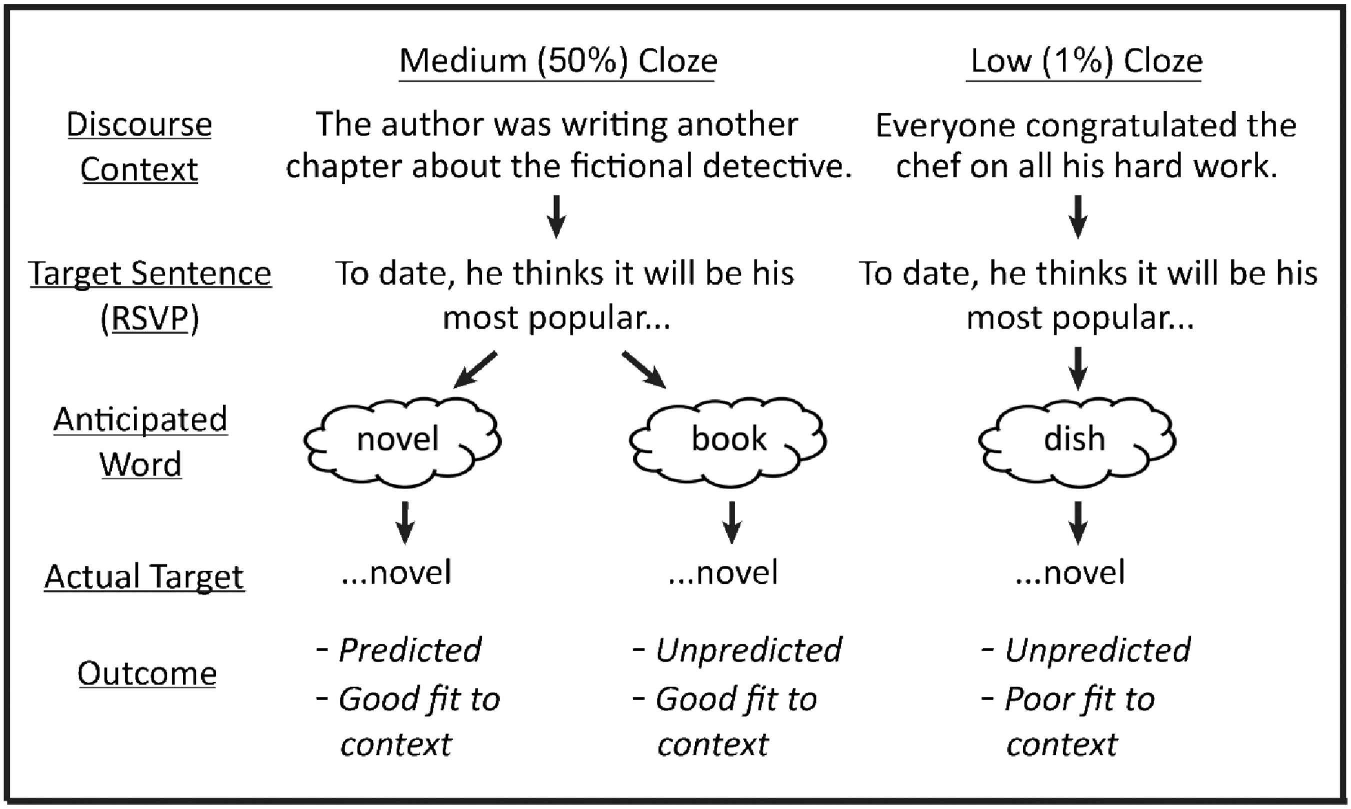

Figure 1.

A summary of the current experimental paradigm. Participants were instructed to actively predict the final word of each discourse and to respond after each trial whether their prediction was correct. ERPs time-locked to the onset of the final critical word (e.g. "novel"), were averaged offline as a function of prediction accuracy (*predicted* vs. *unpredicted*) and contextual support (*medium* vs. *low-cloze*).

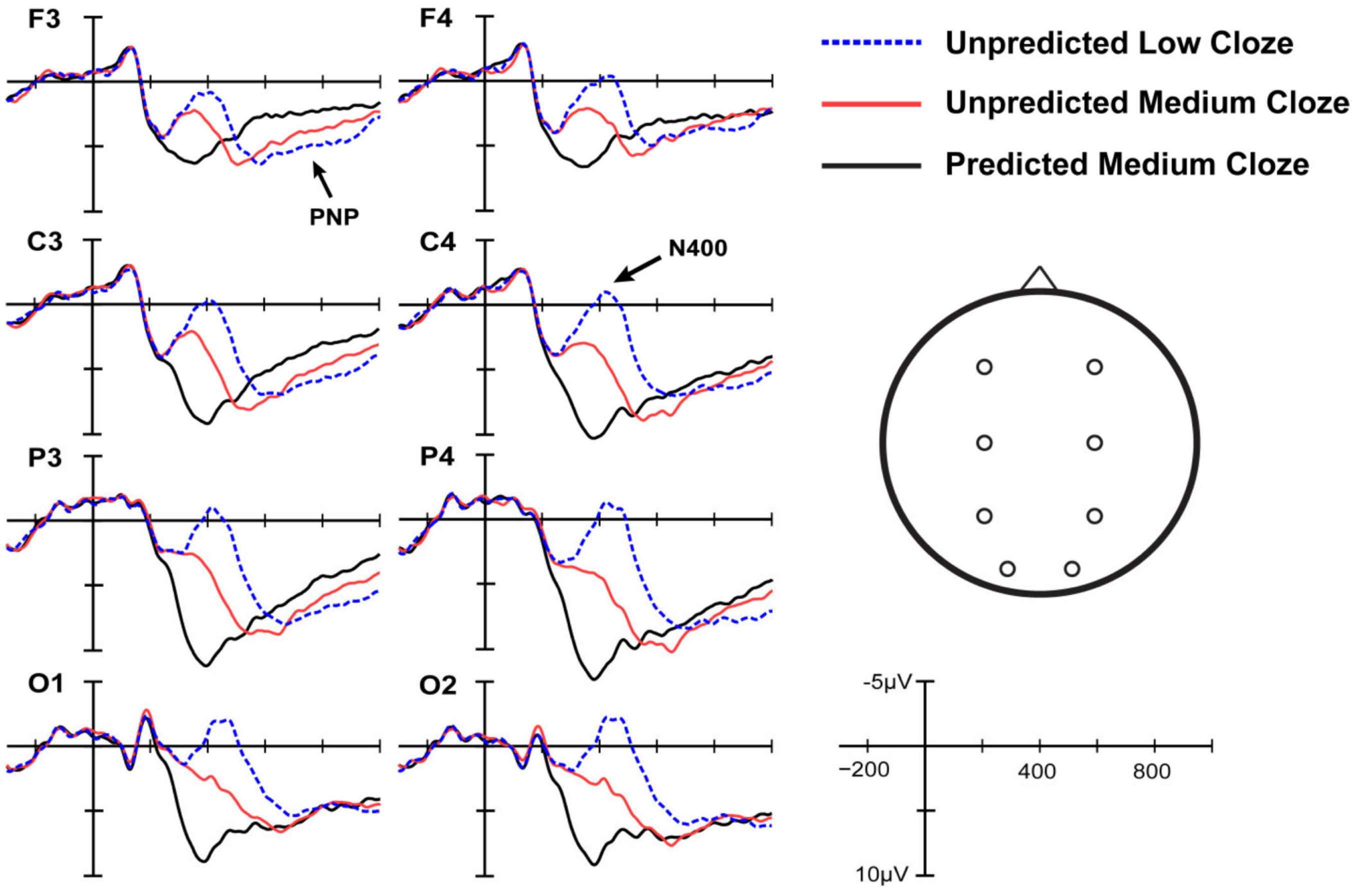

Figure 2.

Averaged event-related potentials for passage-final critical words. All ERP figures were low-pass filtered at 25 Hz for presentation purposes. Separate averages are displayed for unpredicted low-cloze items, unpredicted medium-cloze items, and predicted medium-cloze items.

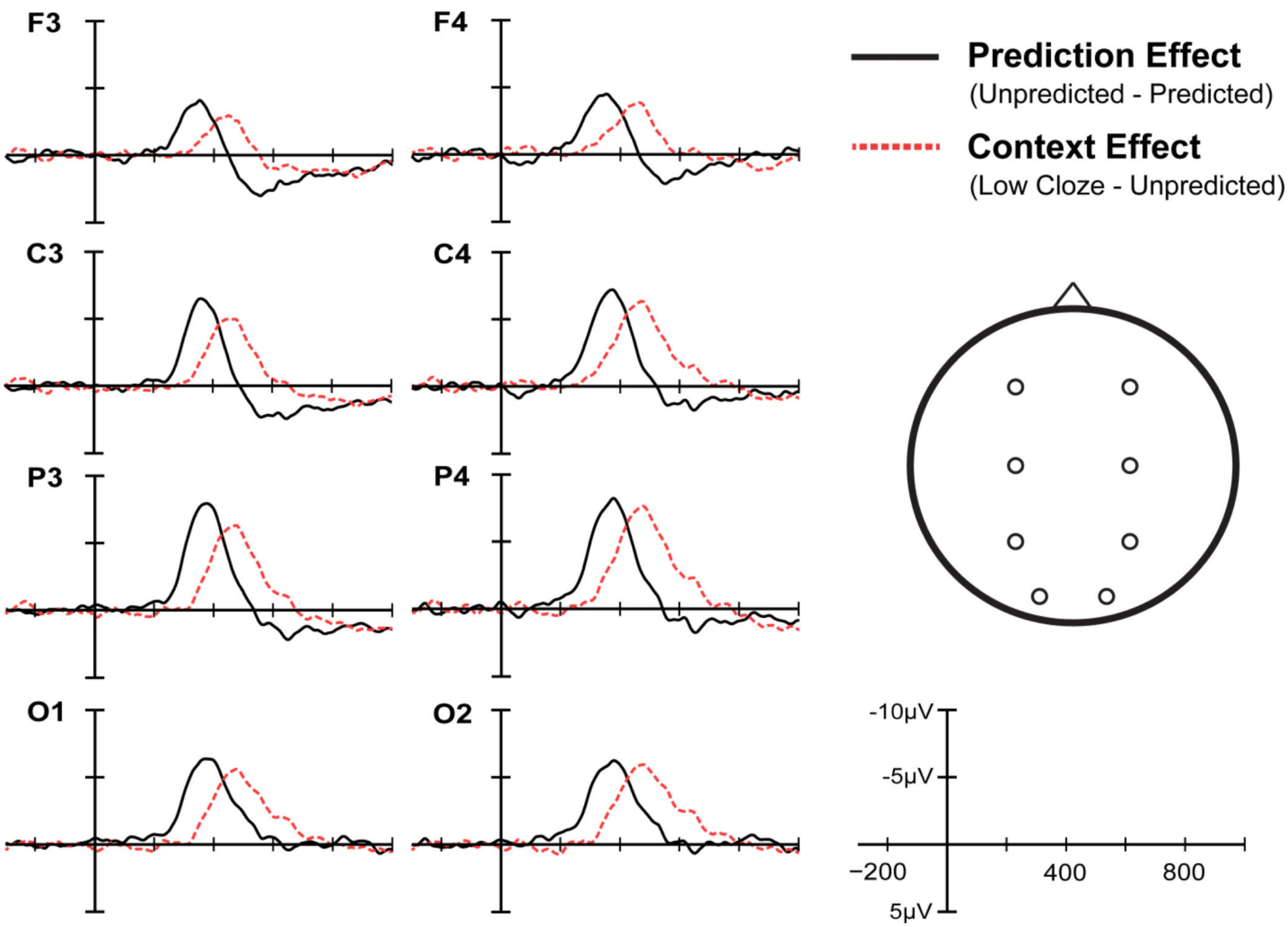

Figure 3.
ERP difference waves representing the effect of prediction accuracy (unpredicted medium-cloze *minus* predicted medium-cloze) and contextual facilitation (unpredicted low-cloze *minus* unpredicted medium-cloze) at the final critical word.

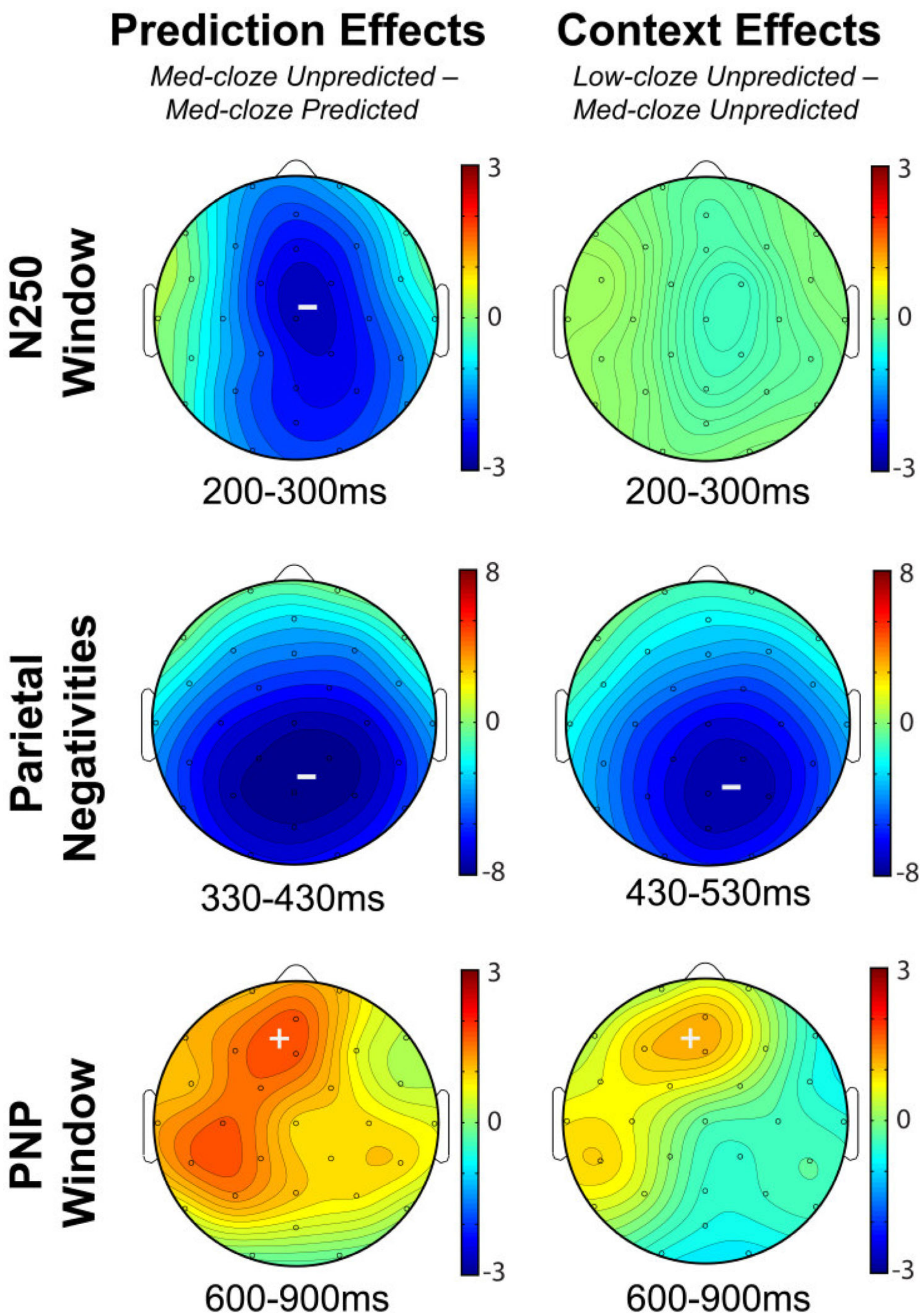

Figure 4.

Topographic distributions of the two grand average difference waves, plotted over time. Note the different voltage scales for each measurement window and the two, separate time windows used for the parietal negativities.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

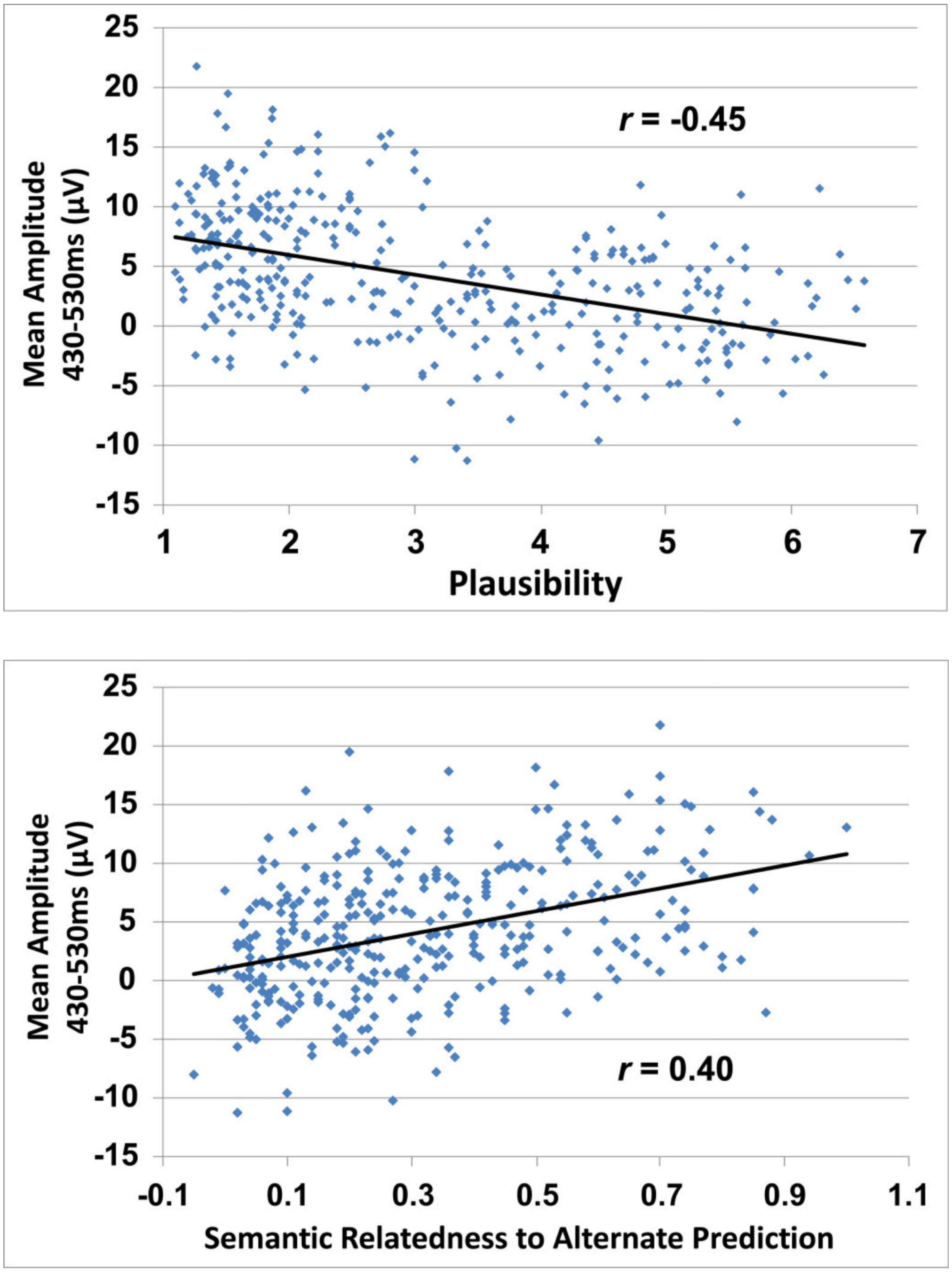

Figure 5.

Scatterplots representing the by-items correlations for unpredicted passage-final words. Values represent mean ERP amplitudes between 430–530ms at central-parietal electrode sites. The top figure shows the relationship between ERP amplitude and passage plausibility (1=makes perfect sense, 7=makes no sense at all). The bottom figure shows the relationship between ERP amplitude and semantic overlap between the critical target and the most-likely alternative completion (calculated using Latent Semantic Analysis). More negative amplitudes were observed for both implausible and semantically unrelated targets.

Table 1

Mean Voltage (430–530ms) at Central-Parietal Electrodes

Source	**Estimate (*b*)**	**Std. Error**	***t*(1,354)**
Intercept	6.35	1.03	6.17 **
Plausibility	−1.24	0.19	−6.55 **
Semantic Overlap	6.41	1.32	4.86 **
Lexical Association	−1.25	2.07	−0.61

$F(3,356) = 39.55$, Adjusted $R^2 = 0.25$

** $p < 0.001$

Table 2

Mean Voltage (600–900ms) at Frontal Electrodes

Source	**Estimate (*b*)**	**Std. Error**	***t*(1,354)**
Intercept	1.89	0.79	2.39 *
Plausibility	0.52	0.15	3.61 **
Semantic Overlap	0.60	1.01	0.59
Lexical Association	−1.28	1.58	−0.81

$F(3,356) = 5.13$, Adjusted $R^2 = 0.033$

* $p < 0.05$,

** $p < 0.001$