THE FATE OF THE UNEXPECTED:
CONSEQUENCES OF MISPREDICTION ON ERP REPETITION PATTERNS

BY

MELINH K. LAI

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Adviser:

Professor Kara D. Federmeier

Abstract

Amid increasing interest in the nature and role of prediction in language comprehension, there remains a gap in our understanding of what happens when predictions are disconfirmed. One possibility is that lingering representations of predictions interfere with those of the unexpected words. Alternatively, it is also possible that violating predictions strengthens the representations of unexpected words – e.g., by drawing attention to them and/or making them more distinctive. Here, the consequences of prediction violations are investigated using the ERP repetition effect. Unexpected but plausible words completed strongly and weakly constraining sentences. Three sentences later the critical word was repeated at the end of a weakly constraining sentence. As a control, the critical word was seen only once in a weakly constraining sentence. In Experiment 1, repeated words elicited a reduced N400 and enhanced LPC, with no effect of initial sentence constraint on the size of the repetition effect in either time window. However, a P2 effect that reached significance only for words initially appearing in a weakly constraining context potentially reflects subtle differences in attention or recognition for words initially appearing in weakly constraining contexts. Experiment 2 used the same items, except for items in the control condition, and added strongly constraining filler sentences with expected endings in order to further promote prediction during reading. Once again, for repeated words there was no effect of initial sentence constraint at either the N400 or the LPC. Additionally, the P2 effect observed in Experiment 1 was not seen in the results for Experiment 2. When considered alongside prior results (Rommers & Federmeier, 2018b, 2018a), these findings suggest that prediction eases processing by instantiating expected words while also potentially reducing encoding or allocation of attention to the input. Thus, the language processor may be well-equipped for the noisiness of language, such that prediction violations may be neither as costly – nor as critical – for language comprehension as has sometimes been assumed.

TABLE OF CONTENTS

Introduction

Historically, the role of prediction in language comprehension has been controversial. A long-standing theoretical view held that, given the typical rapidity and complexity of language input, forming predictions was a task that came at too high a cost for too low a chance of success (e.g., Jackendoff, 2002). To form an online prediction about upcoming input, a person must maintain information from multiple sources within working memory while also correctly identifying how these pieces of information relate to one another. Even when a prediction can successfully be formed, it has a good chance of being incorrect, since sentences can have multiple continuations, intended meanings can be expressed with different lexical items, and the sounds appropriate for any given lexical item may be realized differently by different speakers. On the other hand, despite the potential for variability in any linguistic utterance, all levels of language (including phonology, morphology, syntax, and semantics) follow structured patterns, and language acquisition could essentially be thought of as the development of a fluid understanding of these patterns. In that sense, prediction could theoretically be a very natural phenomenon to arise from the top-down application of linguistic knowledge during online processing. Other arguments in favor of prediction suggest that being able to anticipate upcoming input may ease the burden of processing later on, while also noting that the assumption that incorrect predictions incur costs may not necessarily be accurate (Federmeier, 2007; Van Petten & Luka, 2012)

Indeed, as will be reviewed next, a host of behavioral and neurobiological evidence has emerged supporting the idea that prediction plays a role in language processing. However, the mechanisms of prediction – and its impact on processing over time – remain incompletely characterized. Whereas a growing body of work has examined the in-the-moment effects of predicting and having those predictions confirmed or disconfirmed, little research has addressed the possible downstream consequences of prediction. The goal of the present study is to try to help fill that gap, examining the extent to which predictive processing can have long lasting consequences and providing new insight into the mechanisms at work in making and assessing predictions.

Evidence for predictive processing from behavioral and eye-tracking studies

One line of evidence supportive of a predictive account of language processing comes from studies utilizing eye-tracking in a visual world paradigm (Cooper, 1974; Tanenhaus et al., 1995). In these studies, participants look at different elements of a visual scene while listening to a sentence as a high-speed camera gathers measurement of their saccades and fixation durations, which researchers can then use as indices of referential processing. Altmann & Kamide (1999) used this method to demonstrate that listeners could use information from the verb in a sentence to restrict the referential domain within the visual scene and then correctly anticipate an upcoming word. In their experiment, participants viewed a scene depicting, for example, a boy surrounded by various objects, only one of which was edible (e.g., a cake). While viewing this scene, participants listened to sentences that contained either the verb *eat* or the verb *move*, such as *The boy will eat/move the cake*. Eye-tracking data showed that after hearing *eat* (compared to *move*), participants directed their saccades towards the cake, presumably in anticipation of a patient that was edible. Critically, these anticipatory saccades began before participants heard the end of the determiner (and therefore before they heard the noun), thereby providing a clear instance of contextual facilitation before encountering the critical word *cake*. Other studies using the visual world paradigm have demonstrated similar anticipatory fixations that are driven by other types of linguistic information, such as syntax (Arai & Keller, 2013; Kamide, 2012;

Sussman & Sedivy, 2003), semantic information including thematic roles and concreteness (Duñabeitia et al., 2009; Kamide et al., 2003), and phonological information (Allopenna et al., 1998). The results of these studies thus demonstrate that listeners are able to use information from the input in conjunction with their own linguistic knowledge to correctly direct their gaze to a visual referent in anticipation of an upcoming word. It should be noted, however, that the relationship between processing of linguistic and visual information is still not clear, so that drawing straightforward conclusions from visual world data can often be difficult (Huettig et al., 2011). Therefore, findings from other experimental paradigms provide critical, corroborating evidence about the role of prediction in language processing.

Studies of sentence reading also provide measures that support the hypothesis that language processing involves prediction. Many of these studies use cloze probability as a metric for the predictability of their items (Taylor, 1953). In a cloze norming task, a participant group is given a set of sentence frames and asked to provide the most likely final word. The percentage of times a particular word is given as a completion for a sentence frame constitutes the "cloze probability" of that word in that context. Sentences for which there is relatively high convergence on a particular ending (typically where the cloze probability of the most expected item is above about 60%) are considered highly constraining whereas sentences that do not lead to a strong preference for any one completion – and that, therefore, may be compatible with a more variable set of completions – are termed weakly constraining. Researchers can use cloze probability and sentential constraint to study contextual influences on lexico-semantic processing by comparing how participants read strongly or weakly constraining sentences ending with words of different cloze values.

A wealth of data from behavioral reading studies shows that words are easier to process when they are preceded by congruous sentential contexts, suggesting that the context helps to construct message-level interpretations, which then ease processing of an upcoming word. For instance, numerous studies have shown that response times while reading aloud or during lexical decision tasks (in which participants indicate whether a stimulus is a real word or a pseudoword) are faster when the critical word is preceded by a congruous context (Duffy et al., 1989; Fischler & Bloom, 1979; Forster, 1981; Kleiman, 1980; Schuberth et al., 1981; Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985; Stanovich & West, 1981, 1983; West & Stanovich, 1978). Similar results have been found in patterns of natural reading behavior measured through eye-tracking, wherein congruous words in a highly constraining context are fixated on for less time and skipped more frequently (typically an indication of facilitated processing) than less constrained words (Ehrlich & Rayner, 1981; McDonald & Shillcock, 2003b, 2003a; Rayner et al., 2001; Rayner & Well, 1996). These reading studies thus show that language processing is influenced by contextual constraint and that processing of new information is facilitated when it is consistent with the prior context.

Although this body of evidence is consistent with claims that context can be used predictively, such studies do not necessarily provide direct evidence for predictive processing during comprehension. Other mechanisms, including facilitated integration of incoming information, can also explain the contextual effects on response times in lexical decision tasks and reading aloud (Moss & Marslen-Wilson, 1993; Schwanenflugel & Shoben, 1985; Traxler & Foss, 2000; Tyler & Wessels, 1983). On such an account, initial recognition of the critical word subsequently leads to facilitated integration of the word with information in working memory, which then leads to faster response times in the behavioral task. Integration accounts differ from prediction accounts in that comprehenders do not form an expectation for a particular word or concept, and the facilitative effects of the context on comprehension occur after a word is recognized. Another concern is that research using behavioral tasks may carry some inherent demand characteristics by simply requiring a response from participants, which could encourage prediction in a way that does not accurately reflect what happens during "natural" language processing. An ideal methodology for studying prediction would therefore employ experimental paradigms that specifically isolate predictive processing from other sources of facilitation, using methods that track comprehension with high temporal accuracy while also avoiding the intrinsic demand characteristics of behavioral methods.

Predictive processing research with EEG/ERPs

Use of electroencephalography (EEG) as a method to study language processing and prediction can address or circumvent many of the issues with behavioral methods. In EEG studies, electrodes distributed over the scalp of a participant measure electrical brain activity generated by post synaptic potentials from cortical pyramidal cells (Nunez & Srinivasan, 2006). Most of the insights into language processing from EEG studies comes from the analysis of event related potentials (ERPs), in which researchers take an average of many EEG trials that have been time-locked to the onset of a stimulus (Luck, 2014). EEG and ERPs provide language researchers with a direct neurological measure of language processing that does not rely on a behavioral response. An additional benefit of EEG data is that it is collected with high temporal precision (as fast as every millisecond), which allows researchers to probe the effects of their experimental manipulations almost immediately after they are presented to participants. Finally, EEG records neural activity continuously throughout the experiment, thereby providing data on the neural responses to every word in a sentence and not just a single, critical word. This then allows researchers to look for effects that occur even before the presentation of a critical word, which has crucially helped to differentiate between prediction and integration accounts of contextual facilitation.

ERP researchers take advantage of several ERP components – well-studied features of the ERP wave form linked to specific functional/neural processes – to study language processing, the most well-known of which may be the N400, a negative going wave form with central-posterior topography that peaks around 400 ms after presentation of a stimulus (Kutas & Federmeier, 2011). Although the N400 was first discovered and studied in the written language domain and interpreted as an electrophysiological signature of the processing of semantic anomalies (Kutas & Hillyard, 1980), subsequent studies have shown N400 effects with slight variations (usually in topography) in many different domains, including picture processing (Ganis et al., 1996), face processing (Olivares et al. 1999), auditory language comprehension (McCallum et al., and sign language comprehension (Kutas et al., 1987). The N400 has also been found to be sensitive to a wide range of manipulations, including not only fit to a sentence context, but also repetition and priming (Rugg, 1985; Van Petten et al., 1991). Today the N400 is understood as a general neurophysiological response to complex, potentially meaningful stimuli that may reflect initial contact between that stimulus and information associated with it in long-term semantic memory (e.g., Laszlo & Federmeier, 2011).

The N400 has proven to be a very useful measure for studies of predictive processing in language – and context effects more generally. There is a strong inverse correlation between N400 amplitude and cloze probability, such that words with low cloze values elicit larger N400 amplitudes (Kutas & Hillyard, 1984; many others). More specific to predictive processing, N400 amplitude shows sensitivity to the degree of relatedness between a word that is seen/heard and the one that is expected in that position of a sentence. For example, (Federmeier & Kutas, 1999a) had participants read sentence pairs, in which the first sentence provided a general context while the second sentence ended with an expected word, a semantically related but unexpected word, or a semantically unrelated and unexpected word. An example of a sentence pair from their experiment would be sentences such as *They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of palms/pines/tulips*. They found that while both types of unexpected endings elicited larger N400s than the expected ending (*palms*), the N400s of the semantically related endings (*pines*) were smaller than those of the unrelated endings (*tulips*), despite matched cloze probability and plausibility. Another noteworthy aspect of their findings is that the strength of facilitation for the related items depended on the cloze values for the *expected* words, thereby further demonstrating that this effect was driven by the strength of the prediction for a word that was never actually presented. Similar effects have also been obtained for words that are orthographically related to an expected, but unseen, sentence completion (Laszlo & Federmeier, 2009).

ERP modulations by contextual constraint have also been found on adjectives and articles preceding an expected, critical noun, as a function of whether or not features of the modifier are consistent with properties of the expected – but not yet seen – noun (DeLong et al., 2005; Van Berkum et al., 2005; Wicha et al., 2004). For example, DeLong and colleagues (2005) demonstrated with sentences such as *The day was breezy so the boy went outside to fly a kite/an airplane* that N400 amplitudes at the article preceding low cloze content words were higher than for articles preceding expected content words. These and similar findings have been among the strongest pieces of evidence favoring prediction over integration because they demonstrate contextual influences during the processing of semantically impoverished function words *before* the actual critical word. Although the N400 should not be mistaken as a direct index of prediction, as it appears and is modulated even when predictive processing is diminished/eliminated (Federmeier & Kutas, 1999b; Federmeier et al., 2002), its sensitivity to contextual effects, in conjunction with the right kind of experimental designs, has made it an extremely useful component in the investigation of predictions in language processing.

Another ERP component that has been linked to predictive processing in language is a frontal positivity that temporally follows the N400. Wicha and colleagues (2004) as well as Van Berkum and colleagues (2005) found a frontally distributed positivity following the N400s of the articles preceding unexpected nouns in Spanish and Dutch, respectively. Federmeier and colleagues (2007) also found this post-N400 frontal positivity for unexpected, but plausible, words in strongly constraining English sentences; since then, this effect has been found in a number of other studies as well (DeLong et al., 2014; DeLong et al., 2011; Otten & Van Berkum, 2008; Thornhill & Van Petten, 2012). While this frontal positivity is not yet as well characterized as the N400, its pattern of functional sensitivity, as a response to plausible words that violate strong expectations, has led to suggestions that it might index a processing "cost" associated with disconfirmed predictions, such as suppression of the previously predicted word (DeLong et al., 2014; Otten & Van Berkum, 2008). In support of this, Brothers and colleagues (2015) compared ERPs of medium constraint sentences (with endings of cloze values at approximately 50%) that ended with high cloze or low cloze words, while also collecting participants' responses about whether or not their predictions regarding the final word were correct. Not only did the authors detect a larger frontal positivity for unpredictable items than for predictable items, they also found that the frontal positivity was additionally modulated by participants' explicit detection of a prediction violation. These findings suggest that the post-N400 frontal positivity may be an indication of revision of expectations in relation to the established context, which thus opens the door to investigations of the downstream consequences of valid and invalid predictions.

Critically, the frontal positivity is distinguishable in function and topography from other post-N400 positive effects that have been reported in language (e.g., Van Petten & Luka, 2012). The P600 is a slow, positive-deflection with a centroparietal distribution that is seen around or after 600 ms after stimulus onset, although P600 peaks are often not as distinct as those for the N400 and the timing of this effect is considerably more variable. Originally, the P600 was characterized in the context of syntactic anomalies (Osterhout & Holcomb, 1992), although research from recent years has identified a “semantic P600” that is elicited when words have implausible semantic roles, such as in the sentence *For breakfast, the eggs would only EAT toast and jam* (Kuperberg et al, 2003, see Kuperberg, 2007 for review). Another positive wave form occurring in roughly the same time window as the P600 is the late positive complex (LPC), which is sometimes (but not always) observed to semantic anomalies (e.g., Kos et al., 2012). Similar, posterior positivities have been well-characterized in the memory literature and linked to retrieval from episodic memory. LPC amplitudes are enhanced for deeply encoded words compared to new words (Rugg et al., 1998), for repeated words that are retrieved during explicit memory tasks (Düzel et al., 1997; Rubin et al., 1999), and for recently encountered words that are spontaneously recognized (Kazmerski & Friedman, 1997; Paller et al., 1995; Van Petten & Senkfor, 1996). Like the frontal positivity, these components are often found to be larger for unexpected words; however, whereas the frontal positivity is specific to words that are plausible, the P600/LPC is elicited even – or perhaps especially (DeLong et al., 2014) – by words that are completely anomalous.

Downstream consequences of prediction

Although the debate over whether or not prediction occurs in language processing seems to now lean in favor of prediction, a new line of inquiry has developed from this dispute regarding what circumstances lead to prediction and what consequences prediction has for processing at multiple levels and time points. Indeed, the results from other lines of work provide some indications that predictions may not be ubiquitous or even entirely necessary. For example, N400 data have revealed that older adults show reduced context effects compared to the college-aged adults that are typically tested in psycholinguistic research (Federmeier et al., 2002; Payne & Federmeier, 2018; Wlotko & Federmeier, 2012; Wlotko et al., 2010). Older adults as a group also do not exhibit the post-N400 frontal positivity for unexpected, plausible words in strongly constraining sentences (Wlotko et al., 2012), which has been interpreted as an indication that they may not use context information predictively (Federmeier et al., 2010). Second language (L2) speakers also do not always exhibit the same ERP patterns suggestive of prediction as L1 speakers, with some studies demonstrating a lack of a frontal positivity when comprehending in their L2 (Moreno & Kutas, 2005) or a lack of an N400 effect at articles preceding unexpected words (Martin et al., 2013), while other studies have found such effects when the L2 and L1 are more closely related (Foucart et al., 2014). Finally, even among the native-speaking college-aged population, use of prediction does not seem to be ubiquitous, but, instead, reflects mechanisms that are characteristic of how the left cerebral hemisphere – but not the right hemisphere – processes language. For example, Federmeier & Kutas (1999b) lateralized the presentation of their expected, unexpected but related, and unexpected and unrelated endings to strongly constraining sentences and found that differences between the two unexpected conditions only emerged in left hemisphere-biased processing of stimuli presented in the right visual field. The right hemisphere, instead, appreciated the difference between the expected versus the unexpected endings, but not through predictive processing.

The results from these ERP studies thus indicate that the impacts of prediction unfold in multifaceted ways over time, which could make understanding how prediction affects behavior much less straightforward than originally believed. Indeed, across studies and methodologies, there are notable inconsistencies in the apparent consequences of encountering a word that is different from what might be expected from the context. For example, in a large scale natural reading study using eye-tracking, (Luke & Christianson, 2016) presented participants with passages from a variety of corpora and found that, replicating past work, measures of gaze were slowed for low cloze probability words. However, when participants saw an unexpected but plausible word where a more predictable competitor was available, processing times for the unexpected word actually appeared to be facilitated, rather than slowed. Corresponding findings were obtained in a more controlled experimental design in which both the sentence constraint of the experimental sentences as well as the cloze probability of the target words were manipulated (Frisson et al., 2017). The authors of both studies took their findings as evidence of there being no costs associated with misprediction. This stands in contrast with self-paced reading findings that have found selective slowing to prediction violations, in the same conditions that also cause the frontal positivity (Ng et al., 2017; Payne & Federmeier, 2017). Thus, whether the processing reflected in the frontal positivity is related to behavior – and, if so, how – remains unclear.

Interestingly, Frisson and colleagues (2017) also found that unexpected words that were semantically related to the expected word in a constraining context incurred facilitated reading times in late measures. In this case, the behavioral results seem to line up well with the ERP findings of Federmeier & Kutas (1999a), who saw more facilitation on the N400 for a related (but, in this case, implausible) word when the constraint on the expected word was strong. The eye-tracking findings, however, are again at odds with findings from other behavioral measures. Results from lexical decision tasks have shown that facilitation of unexpected but related endings occurs in sentences with a weakly constraining context, but not if the context is strongly constraining (Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985). Schwanenflugel and colleagues posited that strongly constraining sentences have the effect of narrowing the scope of contextual facilitation so severely that not even processing of semantically related words is facilitated.

Prediction-making thus appears to be a multifaceted process with diverse consequences for the recruitment of neural networks and for behavior – a process that may not always be necessary or even desirable. On the one hand, there is some evidence to suggest that generating predictions and accommodating prediction violations may be taxing and/or incur processing costs. In that case, as hypothesized by early critics of prediction, there may be occasions where it is better for the language processor to hold off on making a prediction. On the other hand, it could also be possible that even encountering prediction errors might be beneficial in the long term, as it is in models of implicit learning (e.g., Chang et al., 2000). A complete account of the prediction mechanism would therefore bridge the gap between models of learning, which posit an important role for prediction violations, and the extant language literature. This requires more information about what happens, not just when a prediction is formed and verified/disconfirmed, but in processing downstream of the prediction as well. A line of work from our lab has thus been looking at the downstream consequences of predicting on explicit and implicit aspects of memory.

A recent set of experiments conducted in our lab has begun to investigate the downstream consequences of predictions for memory, measured both implicitly (Rommers & Federmeier, 2018b, 2018a) and explicitly (Hubbard et al., in press). In an experiment conducted by Rommers and Federmeier (2018b), participants read strongly or weakly constraining sentences ending with a predicted word. Three sentences later, the predicted word reappeared at the end of a weakly constraining sentence. Similar to previous findings that N400 amplitude is reduced when a word is repeated (Rugg, 1985; Van Petten et al., 1991), words that were seen again three sentences after first presentation in a predictive or non-predictive sentence context had reduced N400s in comparison to a baseline condition wherein the critical word was seen for the first time. Critically, however, Rommers and Federmeier (2018b) found smaller N400 repetition effects for words that had previously appeared in a highly predictive context compared to those that had previously appeared in a weakly constraining context. Moreover, following the N400 they found an enhanced positivity (LPC) compared to the baseline condition only in the condition wherein the critical word had first appeared in a weakly constraining sentence, suggesting that when a word was highly predictable at first presentation, repetitions of that word did not yield the same levels of explicit retrieval. These results suggest that one consequence of making accurate predictions in supporting contexts is less thorough processing of the incoming word, leading to reduced processing consequences of (and, by inference, possibly memory for) that same word if repeated. A similar pattern of reduced memory-related ERPs for words that had previously been strongly predictable was also observed in the context of an explicit recognition task (Hubbard et al., in press).

One hypothesis for why predictions might lead to impoverished representations is that the act of predicting serves to partially instantiate the word, such that when it is actually encountered, the system can simply detect the match between its prediction and features of the bottom-up input and concomitantly avoid investing resources in further bottom-up processing and encoding of the input. To test whether prediction does in fact instantiate a representation of the predicted word, in a second experiment, Rommers & Federmeier (2018a) changed the endings of the initial strongly constraining sentences from expected high cloze words to unexpected low cloze words. For example, the sentence *Be careful, the top of the stove is very DIRTY* (cloze = 0.03, according to cloze norms from Federmeier et al. (2007) originally ended with the word *HOT* (cloze = 0.91) in Rommers & Federmeier (2018b). All of the weakly constraining sentences, whether they were the first or second presentation of the critical word, remained unchanged from the first experiment. Thus, in the second experiment, participants had their initial predictions for strongly constraining sentences violated (e.g., they expected *HOT* but were presented with *DIRTY*) but then later, in a different, weakly constraining sentences, they saw the word *HOT*) that they had predicted but had not seen. This allowed Rommers & Federmeier (2018a) to look for lingering activation of the originally predicted word, which they could compare with the repetition effect of seeing a critical word actually appear twice in weakly constraining sentences. Again, a basic repetition effect was seen for items that had initially been encountered in weakly constraining sentences as compared with those that were being seen for the first time. Supporting the hypothesis that predictions are instantiated and can have long-lasting processing consequences, Rommers & Federmeier (2018a) observed what they dubbed a "pseudo-repetition effect" in the N400 time window for the predicted-but-never-seen critical words: N400s for words that were expected but not initially seen were reduced compared to first presentation of words that had never been predicted, although this repetition effect was not as large as that seen for words that had actually been encountered twice. However, prediction-based representations are not identical to those created from bottom-up processing, as the data showed an enhanced late positivity for words that were actually repeated, but not for words that were previously expected but not seen. In their explicit memory task, Hubbard et al., (in press) also found evidence for lingering effects of (mis)predictions, as participants behaviorally false alarmed more and showed dissociable ERP patterns to “lure” words that had been expected but were not seen, compared to matched new words.

Taken together, the results of this set of experiments provide a first look into the downstream consequences of predictive processing. The data suggest that the representations formed during predictive processing are sufficient to create downstream repetition effects at some – but not all – levels of processing and that, more generally, such predictions critically alter processing in two ways. First, it appears that encountering a predicted word results in less thorough processing of the input, such that implicit repetition effects (as indexed by the N400) are diminished and readers do not appear to explicitly retrieve a prior episode of reading the word, as they do if they had previously seen the word without predicting it. It also seems that encountering an unexpected word in the face of a strong prediction is not enough to completely revise a reader’s expectations, as seen in the pseudo-repetition patterns on the N400 and the luring effects observed during explicit recognition. Thus, predictions linger … even when they are disconfirmed.

In turn, these findings raise additional questions about the impact of predicting on the representations that are formed for *other* words in the sentence. In particular, when encountering a word that disconfirms a prediction for a different word, what is the impact of the expectation violation and the lingering representation of the predicted word on the processing of – and memory for – the unexpected word that was actually encountered? One possibility is that having one's prediction violated enhances the resulting representation of the unexpected word, such that downstream memory effects would also be enhanced. Such a result would be consistent with more general findings that memory is often enhanced by distinctiveness and surprise (e.g., Hunt & Elliot, 1980; Konkle et al., 2010; McDaniel & Einstein, 1986). Indeed, Hubbard et al. (in press) found that unexpected (compared to expected) endings elicited larger N1 and LPC amplitudes when they were later encountered on an explicit recognition test, suggesting enhanced attention (N1) and greater levels of explicit recollection (LPC). On the other hand, it is also possible that lingering representations of the originally predicted word could interfere with those of the unexpected words, creating downstream memory costs. Such interference might result if the predicted word acts as a competitor, creating increased demands on lexical selection processes. These results would be consistent with findings that increased lexical competition, such as between the potential meanings of ambiguous words, leads to greater difficulty in processing (Duffy et al., 1988; Rayner & Duffy, 1986) and poorer recall (Klein & Murphy, 2001). As detailed below, the current experiment adapts the design of Rommers and Federmeier to address this question, with the larger aim of learning more about the mechanisms underlying prediction and their impact on language comprehension and memory for verbal material.

Current Study

In two experiments, we adapted the design used by (Rommers & Federmeier, 2018b, 2018a) to examine the downstream processing of words originally encountered as prediction violations. The first experiment used the same initial sentences as (Rommers & Federmeier, 2018a), but then repeated the unexpected, prediction violations in a weakly constraining sentence. In the second experiment, then, we added strongly constraining fillers ending with predicted words (the same ones from Rommers & Federmeier (2018b)). This ensured that participants encountered strong prediction confirmations as frequently as prediction violations, allowing us to test whether effect patterns from Experiment 1 would generalize across different distributions of critical word predictability. At first word presentation, we expected N400 responses graded by cloze probability and enhanced post-N400 frontal positivity to the strong prediction violations compared to that seen for matched words in weakly constraining sentences. Of primary interest in both experiments were N400s to repeated words that had been prediction violations, compared with those to matched repeated words that had previously completed weakly constraining sentences. If encountering prediction violations enhances word processing, we should see a smaller N400 for the second presentation of the word that was initially processed as a prediction violation in a strongly constraining sentence. Additionally, we might observe an enhanced LPC to those words' second presentation, reflecting explicit retrieval of the first appearance of the word. Alternatively, if traces of the originally predicted word continue to linger in memory and cause interference, then first encountering a word as a prediction violation could lead to a diminished repetition effect on the N400 and/or the LPC.

Experiment 1

Methods

Participants

Thirty-five volunteers from the University of Illinois community participated for course credit or payment. All were right-handed native monolingual English speakers with normal or corrected-to-normal vision and no prior history of neurological or psychiatric disorders. Five participants were excluded prior to analysis due to excessive artifacts in the EEG signal (30% of trials or more rejected after artifact detection), leaving a total of 30 participants (16 women and 14 men; average age 20 years, range 18-25 years).

Materials

Sentence stimuli consisted of 123 sentence triplets. Each triplet consisted of one strongly constraining sentence from Federmeier et al. (2007) and two weakly constraining sentence frames. All sentences in a triplet ended with the same word, which was an unexpected ending not only for the weakly constraining frames but also for the strongly constraining sentence (cloze value equal to or near 0). Ending type and position in the repetition sequence led to five sentence types:

(1) Prediction Violation: strongly constraining sentence frames, originally from Federmeier et al. (2007), which ended with an unexpected, but plausible, word

(2) Previously Violation: weakly constraining sentence frames that appeared 3 sentences after a prediction violation sentence and ended with the same word

(3) Unpredictable: weakly constraining sentence frames, adapted from Rommers & Federmeier (2018a, 2018b), which ended with the same unexpected words as in (1)

(4) Previously Unpredictable: weakly constraining sentence frames that followed 3 sentences after an unpredictable sentence and ended with the same word

(5) Not previously seen: weakly constraining sentences whose final critical word had not appeared previously in the block

The Previously Violation, Previously Unpredictable, and Not previously seen sentence types all featured the same critical sentence but differed in terms of the sentences preceding it. Table 1 shows examples, including two intervening filler sentences between the first and second presentation.

Table 1

	Previously Violation
First presentation (SC)	Cats love to be scratched behind the collar. (*expected word = ears*).
Filler	The mother of the tall guard had the same accent.
Filler	The lawyer feared that his client was guilty.
Critical sentence (SCSC)	In the afternoon we started looking for the collar.
	Previously Unpredictable
First presentation (WC)	What caught their attention was something on his collar.
Filler	The mother of the tall guard had the same accent.
Filler	The lawyer feared that his client was guilty.
Critical sentence (WCWC)	In the afternoon we started looking for the collar.
	Not Previously Seen
Filler	The mother of the tall guard had the same accent.
Filler	The lawyer feared that his client was guilty.
Critical sentence (1WC)	In the afternoon we started looking for the collar.

Note. Critical words are underlined. The critical sentence was always weakly constraining, but the conditions differed in terms of what participants had previously seen. If the critical word had previously been seen, it had been shown either in a strongly constraining sentence, where it constituted a prediction violation, or in a weakly constraining sentence.

The sentences in the current study were modified from Rommers & Federmeier (2018b, 2018a) by exchanging weakly constraining sentence frames such that the final critical word was an unexpected but still plausible ending[1]. The degree of constraint of a sentence frame was operationalized as the cloze probability of its most frequently provided completion. As described in more detail in Rommers & Federmeier (2018b, 2018a), sentences were selected based on cloze probability norms collected from native English speakers on Amazon Mechanical Turk (www.mturk.com). In the norming task, participants were told to complete each sentence frame with "the word they would generally expect to find completing the sentence fragment".

In the sentences used for the first presentation, the cloze probability of the critical word was nearly equal in the Prediction Violation sentences (0.002 ± 0.013, range 0.00-0.10) and the Unexpected sentences (0.001 ± 0.007, range 0-0.05), whereas the constraint of the two sentence types was different (cloze probability of the most frequent completion in the strongly constraining Prediction Violation sentences was 0.86 ± 0.13, range 0.45-1.00, whereas in the weakly constraining Unexpected sentences it was 0.19 ± 0.08, range 0.05-0.35). Sentence length was matched (Prediction Violation: 10.02 ± 3.96 words, range 4-21; Unexpected: 10.02 ± 3.95 words, range 4-21). In the critical sentences, the final word had a cloze probability of 0.0008 ± 0.006 (range 0-0.05) and the cloze probability of the most frequent completion was 0.18 ± 0.08 (range 0.05-0.35). The average length of the critical sentences was 8.08 ± 2.23 words (range 4-17 words). Critical words were rotated across conditions, so visual input and all lexical variables were matched.

Three counterbalanced lists were created such that participants would see the critical word from each sentence triplet in only one condition. Each list included an additional 82 fillers, which ensured that over 70% of the sentence endings in a list did not constitute a repetition. The

[1] Several sentences from Rommers & Federmeier could not be arranged such that they ended plausibly. Thus, nine sentences from Rommers & Federmeier's full item list were included in the experiment to maintain the length of the blocks but were not included in analysis.

cloze probability of the sentence endings in fillers was 0.41 on average (range 0.24-0.68). The 282 sentences on each list were distributed across 13 blocks of 21 sentences and one block of 14 sentences[2]. Lists were pseudo-randomized individually for each participant, and critical word repetitions only occurred within a block. On repetitions, critical sentences were separated from the sentence containing the initial presentation by two unrelated sentences; these intervening sentences comprised fillers as well as first presentation sentences or critical sentences belonging to different triplets.

Procedure

Participants were seated in an electrically shielded recording booth approximately 100 cm in front of a 21 inch CRT computer monitor. Stimuli were presented in a white Arial font, size 20, on a black background. Each trial began with a centrally presented crosshair that stayed on the screen for 650 ms followed by a blank screen for 350 ms. Sentences were presented word by word, with each word appearing in the center of the screen for 200 ms with an interstimulus interval of 300 ms. After the final word of each sentence and an interstimulus interval of 1300 ms, three asterisks (* * *) appeared for 2 seconds, indicating that participants were free to blink. Participants took a short rest between blocks.

Participants were asked to minimize blinks, eye movements, and muscle movement while reading for comprehension. They were told that the EEG recording session would be followed by a paper-and-pencil memory test. The recording session began with three practice sentences to introduce participants to the task. Recording time was approximately one hour.

---

[2] Dividing the total number of items across blocks resulted in a shortened final block in which participants saw two items in each experimental condition and four fillers.

Following the recording session, participants completed a paper and pencil recognition test containing 246 words, half of which had been seen as sentence final words and the other half of which had not been seen during the recording session. Unseen words were matched in frequency and length with seen words. Participants were asked to circle all the words they recognized as a sentence-final word from the sentence-reading portion of the experiment.

EEG Recording & Processing

EEG was recorded from 26 evenly spaced silver-silver chloride electrodes mounted on a cap (see Figure 1). Electrodes were referenced online to the left mastoid and re-referenced offline to the average of the left and right mastoids. Additional electrodes were placed on the outer canthus of each eye to monitor for horizontal eye movements and on the left infraorbital ridge to monitor for blinks. All electrode impedances were kept below 5 kΩ. Signals were amplified by a BrainVision BrainAmp DC, with a 16-bit A/D converter, an input impedance of 10 MΩ, a bandpass filter of 0.016-250 Hz, and a sampling rate of 1 kHz. The continuous EEG was high pass filtered offline through a 0.1 Hz filter (two-pass Butterworth with a 12 dB/oct roll-off) and epochs were taken from 750 ms before stimulus onset to 1250 ms after for planned time-frequency analyses that are not reported here. A 200 ms pre-stimulus baseline was subtracted. Raw wave forms were assessed trial-by-trial with artifact thresholds separately calibrated by visual inspection for each subject. Trials were excluded from averaging if they included blinks, movement artifacts, signal drift, blocking, or a horizontal eye movement. On average, a total of 14% of trials (SD = 8%; range across participants: 0%-29%) were marked as artifacts and not included in data analysis. An average of 14% of trials per bin (SD = 0.7%; range across conditions: 13.8%-15.6%) were thus not analyzed.

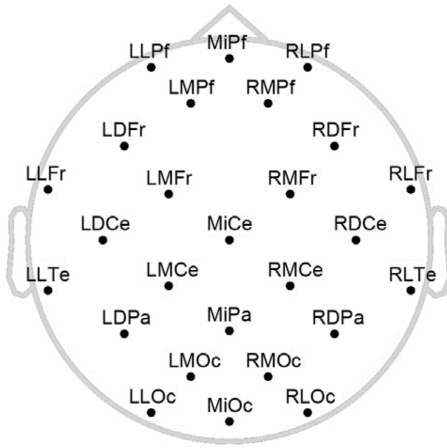


Figure 1. Schematic of the electrode montage with labels.

Event-related potentials. ERPs were formed by averaging in the time domain for each condition and each participant. A 20 Hz low-pass filter was applied (two-pass Butterworth with a 24 dB/oct roll-off). To quantify the N400, mean amplitude measures were taken in an a priori 300-500 ms window, averaged across the LMCe, RMCe, MiCe, MiPa, LDPa, and RDPa channels where N400 amplitude is usually maximal (following Wlotko et al. (2012)). The late frontal positivity mean amplitude at initial word presentations was measured in a 500-800 ms window across 11 frontal channels (MiPf, LLPf, RLPf, LMPf, RMPf, LDFr, RDFr, LMFr, RMFr, LLFr, RLFr; Wlotko et al. (2012)). In the repetition conditions, the late positive complex (LPC) mean amplitude was measured in an a priori 500-800 ms window across the previously mentioned six centroparietal channels, following Rommers & Federmeier (2018a, 2018b). Prior work using this stimulus set (e.g., Wlotko & Federmeier, 2007) has also observed effects on the P2 component , part of the normal visual evoked potential, and visual inspection of the waveforms suggested a possible effect of condition in the P2 time window in the present data as well. Therefore, mean amplitudes were measured in a 190-240 ms window, averaged across the 11 frontal and 6 centroparietal channels.

Results

Memory performance

Participants correctly recognized 37.75% of previously seen words (95% CI [18.8, 27.6]), which was larger than the percentage of false alarms to unseen words (8.73%). This difference was found in all participants who were not already excluded for excessive EEG signal noise and led to a mean *d'* of 1.18 (95% CI [1.023, 1.338]). The larger hit rate to seen words than unseen words across participants suggests that they were paying attention during the EEG session.

The repetition paradigm during sentence reading led to three sub-conditions in the recognition test: Seen Once, Seen First as Unpredictable, and Seen First as Prediction Violation. Trial-level responses (where 1 = judged "seen" and 0 = judged "not seen") were analyzed using a logistic mixed-effects regression model, which simultaneously accounts for participants and items as random effects (Baayen et al., 2008). Along with the fixed effect of Condition, by-item and by-participant random intercepts and by-item random slopes for Condition were entered as predictors (the model failed to converge when by-participant random slopes were included, which would have been the maximal random effect structure warranted by the design; Barr et al., (2013)). There was an effect of Condition, as revealed by a likelihood ratio test of the model relative to an otherwise identical model without the fixed effect of Condition, χ^2 (2) = 54.549, $p < 0.001$. Compared with the Seen Once words (28.7%), Seen First as Prediction Violation (42.7%) were recognized more often by 12.0% (95% CI [8.4, 19.6]), $\beta = 0.697$, $SE = 0.102$, $z = 6.856$, $p < 0.001$. Seen First as Unpredictable words (41.9%) were also recognized more often by 13.2% (95% CI [7.6, 18.8]), $\beta = 0.655$, $SE = 0.096$, $z = 6.862$, $p < 0.001$. There was no evidence for a difference between the Seen First as Prediction Violation and Seen First as Unpredictable conditions (0.8% difference, 95% CI [-3.4, 4.9]), $\beta = 0.041$, $SE = 0.094$, $z = 0.439$, $p = 0.660$.

Thus, repetition enhanced performance, but differences in the predictability of the words during initial presentation did not measurably affect later recognition performance.

Event-related Potentials

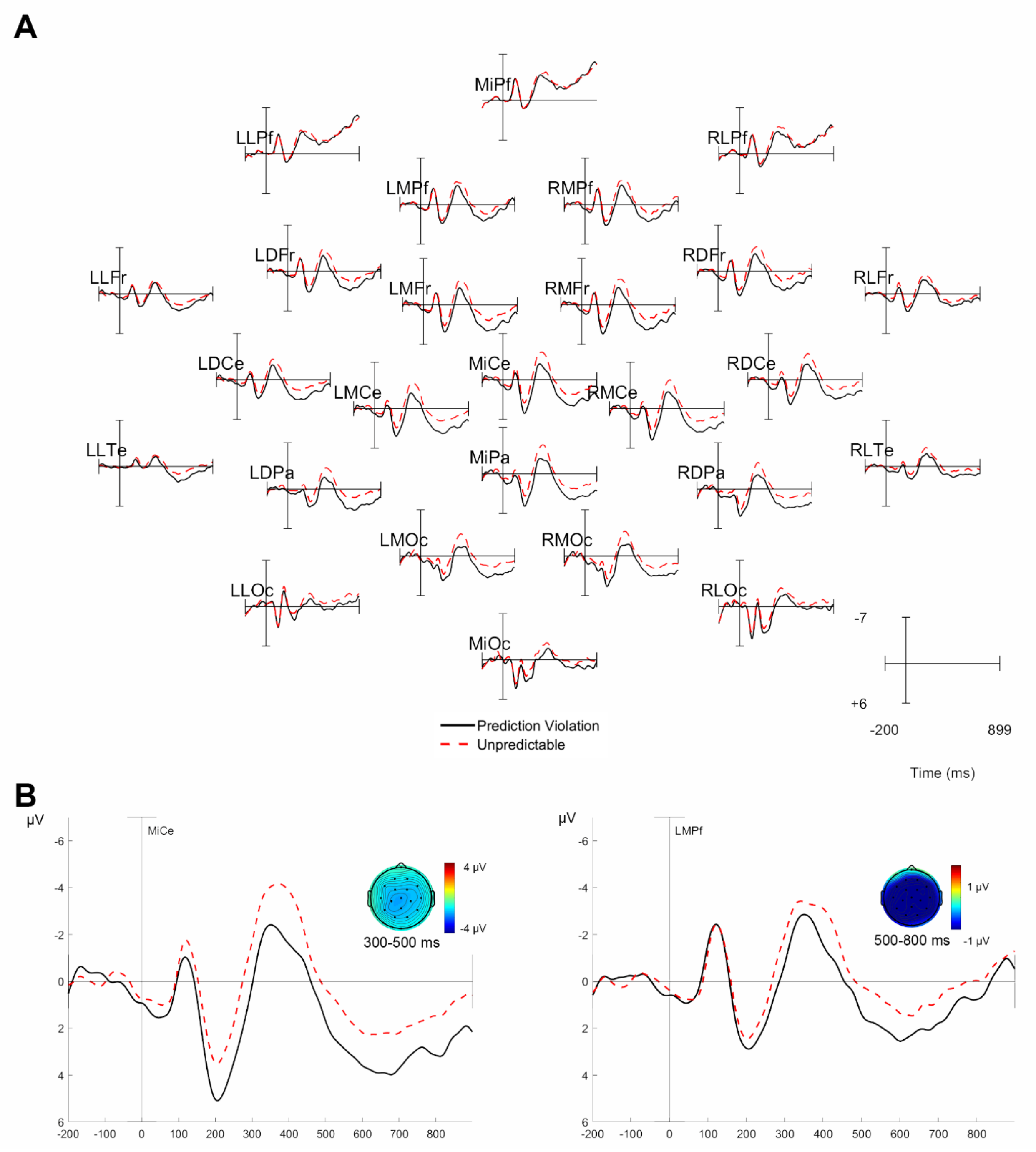


Figure 2. Grand-average ERPs time-locked to words upon initial presentation. Words were prediction violations (seen in strongly constraining contexts) or unpredictable (seen in weakly

Figure 2 (cont.) constraining sentence) A) All scalp electrode sites. B) Close-ups of a central channel (MiCe) and of a frontal channel (LMPf) showing voltage shifts on the P2/N400 and late frontal positivity. Insets show scalp topographies of the difference (Unpredictable – Prediction Violation).

Figure 2 shows the ERPs elicited by the first presentation of the critical words. Following sensory potentials expected for visual presentation (posterior P1, N1, P2; frontal N1, P2), a clear N400 was elicited in both conditions, followed by a late positive-going wave. Consistent with some (Wlotko & Federmeier, 2007) but not other (Rommers & Federmeier, 2018a, 2018b) studies using these same stimuli, condition differences were first observed on the P2, $F(1,29) = 9.308$, $p = 0.00484$, with larger P2 responses to unexpected words that appeared in strongly versus weakly constraining sentences (0.86 µV difference, 95% CI [0.28, 1.44]). Inconsistent with all prior findings, however, these differences between Prediction Violations and Unexpected words continued into the N400 time window: the amplitude of the N400 in response to Prediction Violation words in the present study was attenuated by 1.48 µV (95% CI [0.75, 2.22]) compared with the N400 to Unpredictable words, $F(1,29) = 17.098$, $p = .00028$. As in prior work, responses to Prediction Violations were more positive than to Unexpected words over anterior electrode sites by 0.86 µV (95% CI [0.28, 1.44]), $F(1,29) = 9.308$, $p = 0.00484$). However, this effect was also seen over centroparietal electrode sites, with more positive responses by 1.57 µV (95% CI [0.73, 2.41]) in response to Prediction Violation words compared to Unpredictable words, $F(1,29) = 14.603$, $p = 0.0006$.

As a means of trying to ascertain whether the condition-related differences observed to the first word reflect a single, sustained effect or multiple differences, as well as to compare findings from this study with those of Rommers & Federmeier (2018a), which used the same sentences for their prediction violation condition and almost the same sentences for their

Unpredictable items, except for the critical word (see Footnote 1 in Methods), we overplotted the waveforms from the two experiments; Figure 3.

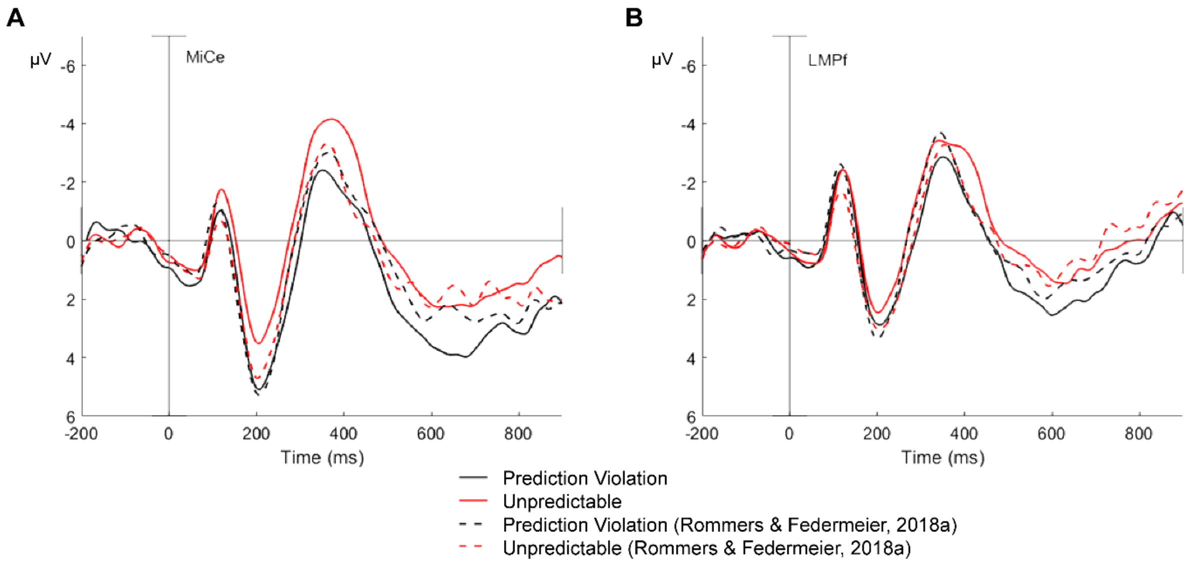


Figure 3. Close-ups of the wave forms of the critical words at first presentation from the current study (solid lines) overlain with the first presentations from Rommers and Federmeier (2018a, dashed lines) at a) a central site (MiCe) and b) a frontal site (LMPf).

This comparison revealed that, compared to Prediction Violations in both studies and Unexpected words in Rommers & Federmeier (2018a), responses to Unexpected words in the present data were more negative over centroparietal sites on both the P2 and the subsequent N400. In the post-N400 window, data from the present study was overall more negative over frontal sites. However, there was also an enhanced positivity, especially over centro-posterior sites to Prediction Violations in the present study compared to those in Rommers & Federmeier (2018a) and compared to Unexpected words in both studies. This qualitative evaluation thus suggests that at least two effects likely contributed to the observed difference in the current data.

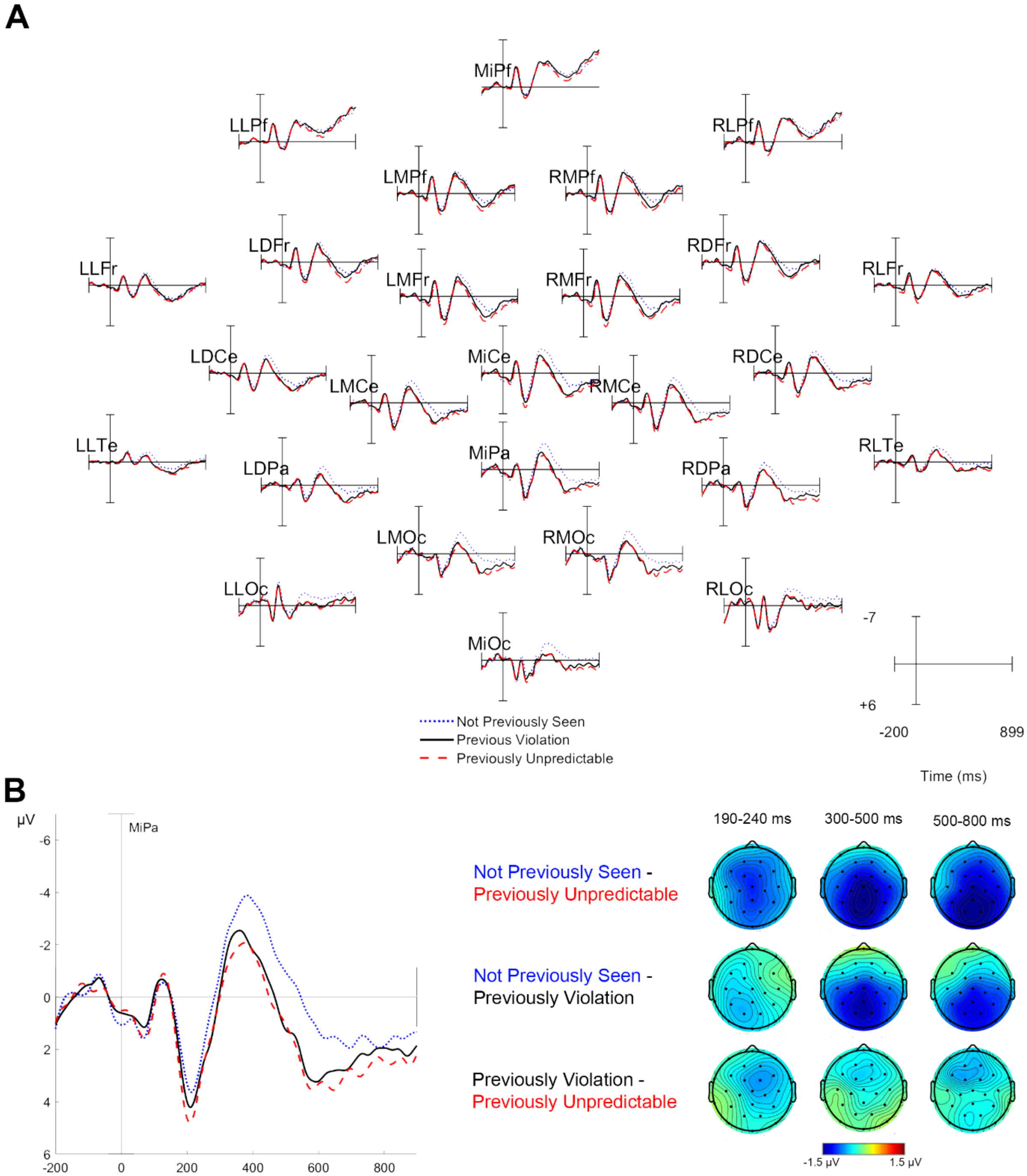


Figure 4. Grand-average ERPs time-locked to critical words in the critical weakly constraining sentences. The words were either repetitions (Previously Violation, Previously Unpredictable) or unseen words presented in the same sentence contexts (Not Previously Seen). A) All scalp electrode sites. B) Close-up of a central channel (MiPa) showing the P2, N400, and LPC, along with scalp topographies of the repetition effect for previously unpredictable words, the repetition effect for previously predictable words, and the effect of prior predictability.

Figure 4 shows the ERPs elicited by the critical words in the critical repetition sentences. Condition-related differences were first apparent on the P2 and continued into the N400 and LPC time windows. P2 amplitude across the frontal and centroparietal channels differed between conditions, $F(2,58) = 3.798$, $p = 0.0028$ (Greenhouse-Geisser corrected, $\varepsilon = 0.997$). Relative to the Not Previously Seen words, P2 amplitude in response to Previously Unpredictable words was more positive by 0.81 µV (95% CI [0.21, 1.40]), $F(1,29) = 7.617$, $p < 0.01$. However, P2 repetition effects were not significant for Previously Violation (compared to Not Previously Seen) words: difference of 0.33 µV (95% CI [-0.29, 0.95]), $F(1,29) = 1.188$, $p = 0.28$. The difference in the size of the P2 repetition effects across the two repetition conditions was marginal, with Previously Unpredictable words more positive by 0.48 µV (95% CI [-0.40, 1.06]) compared with the Previously Violation words, $F(1,29) = 2.74$, $p = 0.110$.

N400 amplitude differed between conditions, $F(2,58) = 17.191$, $p < 0.001$ (Greenhouse-Geisser corrected; $\varepsilon = 0.972$). Relative to the Not Previously Seen words, which elicited the largest N400, the N400 in response to Previously Unpredictable words was attenuated by 1.46 µV (95% CI [0.92, 2.00]), $F(1,29) = 30.829$, $p < 0.001$. The N400 in response to Previously Violation words was attenuated by 1.21 µV (95% CI [0.70, 1.72]), $F(1,29) = 23.715$, $p < 0.001$. Thus, both predictable and unpredictable words elicited repetition effects on the N400 when presented again. Although the repetition effect for Previously Violation words was numerically smaller than that for Previously Unpredictable words by 0.25 µV, this difference was nonsignificant (95% CI [-0.33, 0.84]), $F(1,29) = 0.779$, $p = 0.385$.

The amplitude of the late positive complex (LPC) also differed between conditions, $F(2,58) = 10.216$, $p < 0.001$ (Greenhouse-Geisser corrected, $\varepsilon = 0.950$). Relative to the Not Previously Seen words, LPC amplitude in response to Previously Unpredictable words was more positive by 1.39 μV (95% CI [0.76, 2.02]), F (1,29) = 20.069, p = 0.0001. The LPC in response to Previously Violation words was more positive by 1.05 μV (95% CI [0.45, 1.67]), F (1,29) = 12.806, p = 0.001. The LPC to Previously Unpredictable words did not significantly differ from the Previously Violation words (95% CI [-0.40, 1.06], F (1,29) = 0.900, p = 0.351).

Overall, then, repetition effects were obtained on N400 and LPC, replicating past findings with this paradigm Critically, however, these effects did not significantly differ between Previously Unexpected words and Previously Violations. We also observed repetition effects on the P2, which were significant only for the Previously Unexpected words.

The verbal fluency scores of the semantic categories (animals, first names, and fruits and vegetables) were averaged (pairwise correlations ranged from r = 0.296 to r = 0.692). Confirming previous work linking verbal fluency to N400 patterns, we found an effect of verbal fluency on the N400s elicited at initial presentation, here in the form of participants with better semantic verbal fluency showing a somewhat larger N400 at initial presentation, r = 0.405, p = 0.026. Verbal fluency was not strongly correlated with effects on the post-N400 positivity at initial presentation, or with the N400 or LPC effects at second presentation, rs < 0.11, ps > 0.5.

Experiment 1 Discussion

Participants in Experiment 1 read strongly or weakly constraining sentences with unexpected, but plausible, endings. Three sentences later, the critical sentence-final word appeared again at the end of a weakly constraining sentence. In a control condition, the critical word was seen only once at the end of a weakly constraining sentence. Sentences in both second presentation conditions showed consistent facilitative effects of repetition, with repeated words eliciting a reduced N400 as well as an enhanced LPC in comparison to the control condition. However, there were no differences observed at either the N400 or the LPC for words that had previously appeared as prediction violations versus previously unpredictable words. A similar pattern was seen in results from a paper-and-pencil recognition test of the critical words, where there was a clear benefit of repetition but no effect of prior predictability. Taken together, results thus far suggest that processing prediction-violating words ultimately incurs neither cost nor benefit to reprocessing of the word when it is seen again downstream.

However, there was one difference between the downstream response to previous prediction violations and words that were previously unexpected but not violations of strong expectations: A post-hoc analysis of P2 amplitude across frontal and centroparietal channels found a repetition effect that reached significance only for unpredictable words initially appearing in weakly constraining sentences. P2 effects in the context of sentence processing (Federmeier et al., 2005; Wlotko & Federmeier, 2007) have been argued to reflect a change of state in the processing system in anticipation of, for example, a probable predicted stimulus. Enhanced P2 amplitudes at frontal or frontocentral sites have been reported for repeated words in a number of paradigms (Misra & Holcomb, 2003; Van Petten et al., 1991; Woollams et al., 2008). The sensitivity of the P2 to repeated visual words suggest that the P2 may reflect early procedures of word recognition or lexical access. Under this account, the P2 effect found in this experiment could reflect a difference in how the system recognizes or accesses a repeated word that had originally appeared in a weakly constraining sentence. In a similar vein, although the P2 difference between the Previously Violation and Previously Unpredictable conditions did not reach significance ($p = 0.110$), the intermediate P2 amplitude of the Previously Violation condition might suggest that perceptual and/or attentional resources were taxed during the encoding of unexpected words when they were encountered as strong prediction violations.

Indeed, ERP differences seen on initial presentation provide supporting evidence that processing prediction violations is different than processing unexpected words in the context of weakly constraining sentences that do not permit strong expectations about the sentence-final word. At initial presentation, we observed a reduced P2 response to unexpected words in weakly constraining contexts compared to the prediction violations seen in strongly constraining contexts. Similar patterns of constraint-related P2 effects have been seen for lateralized stimuli (e.g., Federmeier & Kutas, 2002; Federmeier et al., 2005; Wlotko & Federmeier, 2007). The pattern from these previous studies suggests that attentional resources may have been employed differently in strongly versus weakly constraining sentences. However, this P2 effect was not found in either of the prior experiments using this paradigm (Rommers & Federmeier, 2018a, 2018b), and, while this may further suggest an early attention-driven difference in processing unexpected stimuli in prediction-generating contexts, it also warrants continued exploration in order to fully delineate the nature of the effect.

The effect difference on the P2 then continued into the N400 time window. It is difficult to determine if this is a separate effect on the N400 or a lasting influence of the voltage shift elicited in the prior window. N400 differences are not typically obtained to comparisons of unexpected words across constraint conditions when these are matched for cloze probability (e.g., Kutas & Hillyard, 1984; Federmeier et al., 2007 using similar materials) and no N400 differences were observed in Rommers & Federmeier (2018a) for the similar comparison here, lending support to the idea that this effect pattern may reflect continued influence of the P2-related shift. However, it should be noted that the unexpected words in the present study are not the same endings as those in Rommers & Federmeier (2018a) and that measures of cloze probability are essentially at floor for these conditions, making it difficult to know if there are differences in probability or plausibility that were not captured in the norming and could be contributing to the observed difference.

The effect pattern continued further into the late positivity time window. While words in the Prediction Violation condition elicited a more positive-going wave over frontal areas than did the Unpredictable condition, this effect was also evident over central and parietal regions, making it difficult to determine what the overlaps is between this effect and plausible prediction violation effects previously described in the literature (e.g., Federmeier et al., 2007, among others). One explanation for this pattern of distribution may be the occurrence of an additional LPC effect, in tandem with the frontal positivity; this might suggest that participants found some of the prediction violations implausible.

Overall, the pattern of results for second presentation sentences, where the critical word was previously encountered in prediction-violating or weakly constraining contexts, as well as the results from the paper-and-pencil recognition test, suggest that encountering an unexpected as a prediction violation has no benefit or cost over encountering the same word in a weakly constraining sentence for downstream processing. However, given the unexpected effects found at first presentation, we wanted to run a second experiment using the same critical conditions to get a better handle on which aspects of the observed pattern are consistent. Moreover, a disadvantage of the design of Experiment 1 is that all of the strongly constraining sentences ended with unexpected words. Although there were medium constraint fillers that ended predictably, the fact that predictions were so often violated might have tended to discourage participants from engaging in prediction. We did obtain a post-N400 positivity effect that has been previously linked to predictive strategies, but, as described, the distribution of the effect was broader than typical. Thus, we adapted the follow-up experiment to further encourage the use of prediction. In Experiment 2, participants read the same critical stimuli from the Previously Violation and Previously Unpredictable conditions of Experiment 1. However, to ensure that participants saw as many prediction confirmations as prediction violations, we replaced the Not Previously Seen condition with the strongly constraining sentences from Rommers & Federmeier (2018b) all of which ended with the most predicted word. This means we will not have the well-matched control condition for assessing the size of the repetition effect compared to items seen only once. However, basic N400 repetition effects are already well-established, not only in Experiment 1 but also in the prior (Rommers & Federmeier, 2018b, 2018a) studies and in the wider literature. In Experiment 2, the critical comparison continues to be between the previously unpredictable and previously violation conditions, both of which are repeated in the same sentence frame.

Experiment 2

Methods

Participants

Forty-two volunteers from the University of Illinois community participated for course credit. All were right-handed native monolingual English speakers with normal or corrected-to-normal vision and no prior history of neurological or psychiatric disorders. None participated in Experiment 1. Twelve participants were excluded prior to analysis due to excessive artifacts in the EEG signal (30% of trials or more rejected after artifact detection), leaving a total of 30 participants (14 women and 16 men; average age 19.4 years, range 18-22 years).

Materials

Critical sentence stimuli from Experiment 1 (except for the Not Previously Seen condition) were combined with the Predictable and Unpredictable items from Rommers & Federmeier (2018a). Ending type and position in the repetition sequence led to six sentence types:

(1) Prediction Confirmation: strongly constraining sentence frames, originally from Federmeier et al. (2007), which ended with an expected word

(2) Unpredictable with Confirmation Ending: weakly constraining sentence frames from Federmeier et al. (2007), which ended with the same expected word as in (1)

(3) Prediction Violation: unchanged from Experiment 1

(4) Previously Violation: unchanged from Experiment 1

(5) Unpredictable with Violation Ending: sentences from the Unpredictable condition from Experiment 1, unchanged

(6) Previously Unpredictable: unchanged from Experiment 1

The Previously Violation and Previously Unpredictable sentence types once again featured the same critical sentence but differed in the sentences preceding it. Table 2 shows examples, including two intervening filler sentences between the first and second presentation.

Table 2

	Prediction Confirmation
Single presentation	Cats love to be scratched behind the ears.
	Unpredictable with Confirmation Ending
Single presentation	The baby cried when somebody touched his ears.
	Previously Violation
First presentation	Cats love to be scratched behind the collar. (*expected word = ears*).
Filler	The mother of the tall guard had the same accent.
Filler	The lawyer feared that his client was guilty.
Critical sentence	In the afternoon we started looking for the collar.
	Previously Unpredictable
First presentation	What caught their attention was something on his collar.
Filler	The mother of the tall guard had the same accent.
Filler	The lawyer feared that his client was guilty.
Critical sentence	In the afternoon we started looking for the collar.

Three counterbalanced lists were created such that participants would see the critical word in only one condition. As in Experiment 1, the lists included an additional 82 fillers, which ensured that 75% of the sentence endings in a list did not constitute a repetition. The 328 sentences on each list were distributed across 13 blocks of 24 sentences and one block of 16 sentences. Lists were pseudo-randomized individually for each participant, and critical word repetitions only occurred within a block. On repetitions, critical sentences were separated from the sentence containing the initial presentation by two unrelated sentences; these intervening sentences comprised fillers as well as first presentation sentences or critical sentences belonging to different triplets.

Recognition test. The 123 expected endings (which were only seen once) from Rommers & Federmeier (2018b) were added to the recognition test from Experiment 1, along with an additional 52 words that had not appeared in any of the stimuli. Words that appeared as expected or unexpected items in one experimental list may not have appeared as a sentence ending in another list; these words were coded as "new" for the lists in which they did not appear. Twelve words (24 items) that appeared in both Experiment 1 as an unexpected word and in Rommers & Federmeier (2018b) as an expected word were removed from analysis. An additional 7 words were removed for being the plural form of another word on the test (e.g., *keys* and *key*). The final test consisted of 267 words, and for participants from any of the experimental lists, 144 words had been seen as sentence final words and the remaining 123 had not been seen during the recording session.

Procedure

The experiment procedure was the same as Experiment 1. Participants again took the paper and pencil recognition test following the EEG recording session.

EEG Recording & Processing

EEG recording, baselining, and trial exclusion procedures were the same as Experiment 1. On average, a total of 20% of trials (SD = 8%; range across participants: 5%-30%) were marked as artifacts and not included in data analysis. An average of 20% of trials per bin (SD = 1.2%; range across conditions: 19.1%-23.0%) were thus not analyzed.

Results

Memory performance

Participants correctly recognized 32.93% of previously seen words (95% CI [15.7, 23.2]), which was larger than the percentage of false alarms to unseen words (13.43%). This difference was found in all participants who were not already excluded for excessive EEG signal noise and led to a mean *d'* of 0.749 (95% CI [0.615, 0.882]). The larger hit rate to seen words than unseen words across participants suggests that they were paying attention during the EEG session.

The inclusion of items from Rommers & Federmeier (2018b) with the items in the repetition paradigm during sentence reading led to four sub-conditions in the recognition test: Seen Once as Prediction Confirmation, Seen Once as Unpredictable, Seen First as Unpredictable, and Seen First as Prediction Violation. The logistic regression model analyzing trial-level responses was otherwise the same as in Experiment 1. There was an effect of Condition, as revealed by a likelihood ratio test of the model relative to an otherwise identical model without the fixed effect of Condition, χ^2 (3) = 31.586, $p < 0.001$. Compared with the Seen Once as Prediction Confirmation (24.1%) and Seen Once as Unpredictable words (28.3%), Seen First as Prediction Violation (39.9%) were recognized more often by 15.8% (95% CI [11.1, 20.4]), β = 0.780, *SE* = 0.144, z = 5.421, $p < 0.001$, and 11.6% (95% CI [6.9, 16.3]), β = 0.529, *SE* = 0.139, z = 3.808, p = .0001, respectively. Seen First as Unpredictable words (39.2%) were also recognized more often by 15.1% (95% CI [10.4, 20.0]), β = 0.755, *SE* = 0.145, z = 5.218, $p <$ 0.001 and 10.9% (95% CI [7.2, 14.8]), β = 0.504, *SE* = 0.140, z = 6.862, p = .0003, respectively. There was no evidence for a difference between the Seen First as Prediction Violation and Seen First as Unpredictable conditions (0.7% difference, 95% CI [-4.3, 5.5]), β = 0.024, *SE* = 0.105, z = 0.231, p = 0.817. Thus, results from Experiment 1 replicated, with repetition enhancing performance, and differences in the predictability of the words during initial presentation failing to measurably affect later recognition performance.

Event-related Potentials

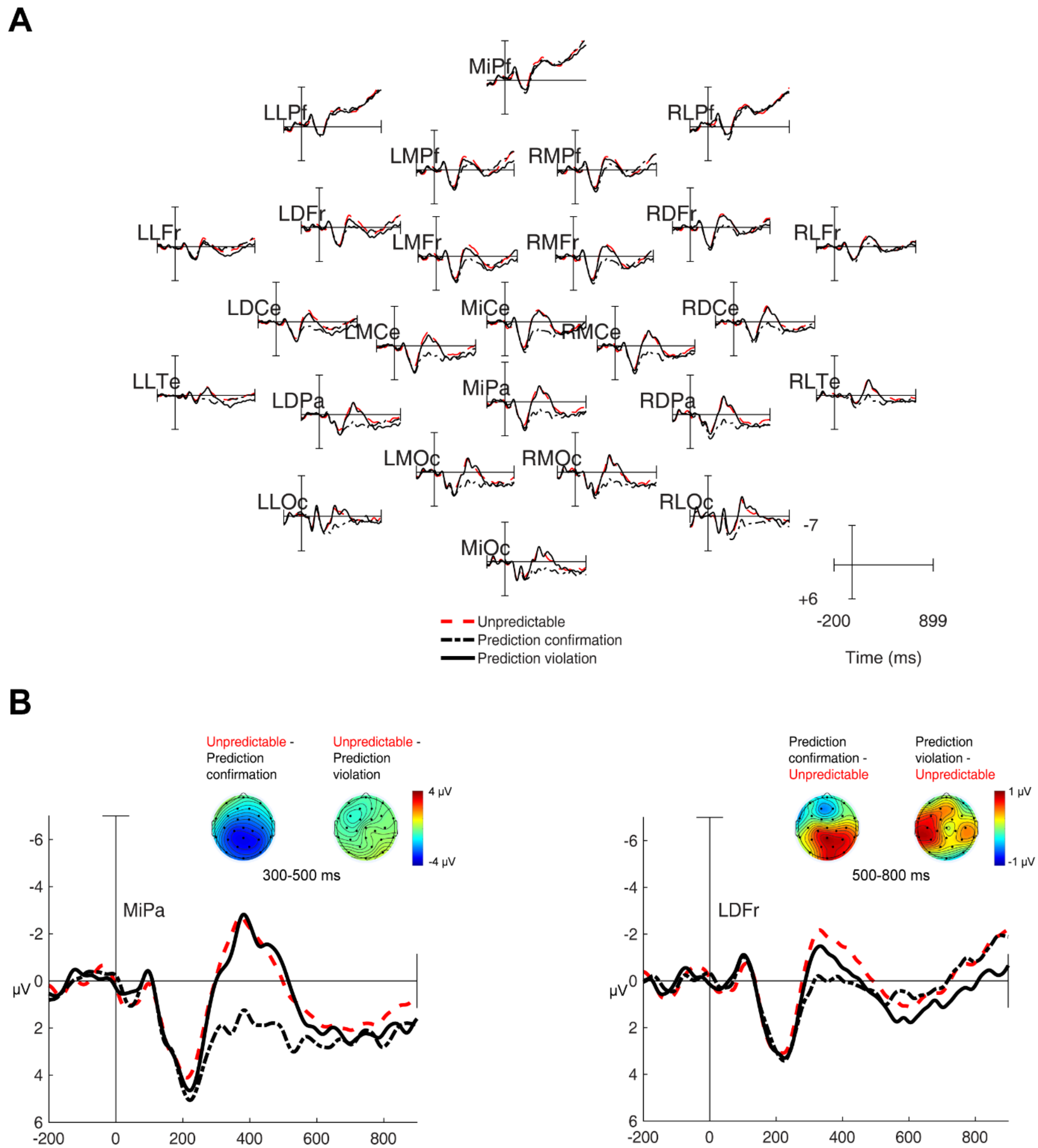


Figure 5. Grand-average ERPs time-locked to words upon initial presentation. Words were prediction violations or prediction confirmations (presented in strongly constraining contexts) or unpredictable (presented in weakly constraining sentence). Negative is plotted up. A) All scalp

Figure 5 (cont.) electrode sites. B) Close-ups of a central channel (MiPa) and of a frontal channel (LDFr) with scalp topographies.

Following sensory potentials expected for visual presentation (posterior P1, N1, P2; frontal N1, P2), a clear N400 was elicited in all conditions, followed by a late positive-going wave. There was no evidence of a difference between the two weakly constraining initial sentences, Unpredictable with Confirmation Ending and the Unpredictable with Violation Ending sentences, on the P2 (difference = 0.018 μV (95% CI [-0.782, 0.746]), $F(1, 29) = 0.002$, $p = 0.961$) or the N400 (difference = 0.167 μV, (95% CI [-0.587, 0.920]), $F(1, 29) = 0.205$, $p = 0.654$). There was a small difference in the late time window, with Unpredictable with Violation Ending sentences eliciting a more positive deflection by about 0.88 μV, (95% CI [0.077, 1.684]), $F(1,29) = 5.028$, $p = 0.033$. To simplify analyses, these conditions were combined into a single Unpredictable condition to be used as a comparison with both Prediction Confirmation and Prediction Violation sentences.

Figure 5 shows the ERPs elicited by the words at the end of Prediction Confirmation sentences and the first presentation of the critical words in Prediction Violation sentences and Unpredictable sentences. No condition differences were observed between the Prediction Confirmation, the Prediction Violation, or the Unpredictable words on the P2, $F(1,29) = 1.350$, $p = 0.0672$ (Greenhouse-Geisser corrected, $\varepsilon = 0.996$). Relative to the Unpredictable words, P2 amplitude in response to Prediction Confirmation words differed by 0.536 μV (95% CI [-1.21, 0.138]), $F(1,29) = 2.65$, $p = 0.11$. P2 effects were likewise not significant for Prediction Violation (compared to Unpredictable) words, with a difference of 0.259 μV (95% CI [-0.94, 0.42]), $F(1,29) = 0.606$, $p = 0.92$. There was no difference in the size of the P2 effects across the Prediction Violation and Prediction Confirmation conditions, difference = 0.277 μV (95% CI [-

0.37, 0.92]), F (1,29) = 0.017, p = 0.39. Thus, different from the pattern in Experiment 1, we found no evidence here that P2 responses were larger to words in strongly constraining contexts.

An effect of condition was seen in the N400, F(2, 58) = 41.185, p < 0.001 (Greenhouse-Geisser corrected, ε = 0.822). Different from the results from Experiment 1 but consistent with the results from Rommers & Federmeier (2018b), N400 amplitudes in response to Prediction Violation words in the present study did not significantly differ from those to Unpredictable words, with a numerical difference of 0.046 µV (95% CI [-0.659, 0.751]), F(1, 29) = 0.0178, p = 0.895. Consistent with canonical findings, including Rommers & Federmeier (2018b), Prediction Confirmation words manifested a greatly reduced N400 amplitude in comparison to Unpredictable words (difference of 2.97 µV (95% CI [2.309, 3.632], F(1,29) = 84.4, p < 0.001). Similarly, N400 amplitude for Prediction Confirmation words was reduced relative to Prediction Violation words by 3.02 µV (95% CI [2.075, 3.958], F(1,29) = 42.9, p < 0.001.

There was not evidence of an overall effect of condition over anterior electrodes, F(2, 58) = 2.467, p = 0.0937 (Greenhouse-Geisser corrected, ε = 0.837). However, pairwise comparisons showed that Prediction Violations were more positive compared to Unpredictable words over anterior electrode sites by 0.756 µV (95% CI [0.17, 1.34]), F (1,29) = 7.023, p = 0.0129). The distribution of this effect over centroparietal electrodes was not found to be significant, with a numerical difference of 0.534 µV [-1.276, 0.208]), F(1,29) = 2.17, p = 0.152, suggesting that the slow potential effect found in Experiment 1 was not apparent in this dataset. This frontal positivity was not seen in the comparison between Prediction Confirmation words and Unpredictable words, with a numerical difference of 0.536 µV (95% CI [-1.210, 0.138], F(1,29) = 7.023, p = 0.520.

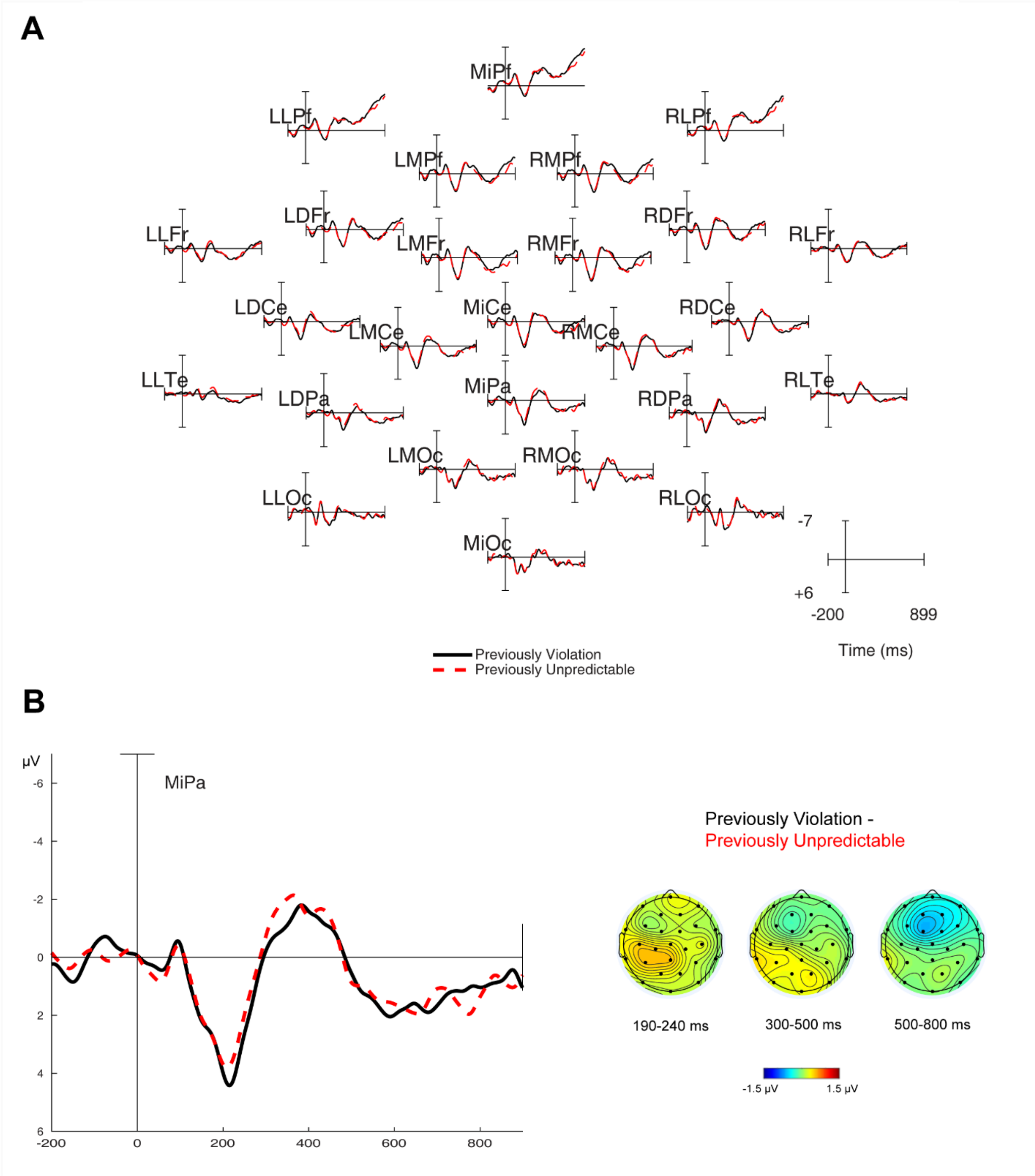


Figure 6. Grand-average ERPs time-locked to critical words in the weakly constraining Previously Violation and Previously Unpredictable sentences. All words were repetitions. A) All scalp electrode sites. B) Close-up of a central channel (MiPa) and scalp topographies depicting a null effect of prior predictability.

Figure 6 shows the ERPs elicited by the repeated critical words. There were no differences between the critical repeated conditions in any time window. There was no difference in the size of the P2 across the two repetition conditions (difference = 0.03 μV (95% CI [-0.60, 0.67]), F (1,29) = 0.017, p = 0.91); recall that there had been a tendency toward a difference between these conditions in Experiment 1. As in Experiment 1, the N400 to Previously Unpredictable words did not significantly differ from the Previously Violation words (95% CI [-0.304, 1.238], F (1,29) = 1.532, p = 0.226). The LPC to Previously Unpredictable words also did not significantly differ from the Previously Violation words (95% CI [-0.84, 0.64], F (1,29) = 0.047, p = 0.830). Overall, then, replicating Experiment 1, there was no evidence for downstream effects of encountering a prediction violation compared to the effect of processing a matched, unpredictable word that completed a weakly constraining sentence frame.

The verbal fluency scores of the semantic categories animals, first names, and fruits and vegetables were averaged (pairwise correlations ranged from r = 0.237 to r = 0.576). In this experiment, we did not find an effect of verbal fluency on the N400s of Prediction Confirmation sentences in comparison to Unpredictable sentences, r = 0.167, p = 0.378. Verbal fluency again was not strongly correlated with the late positivity effect or with the N400 or LPC effects at second presentation, rs < 0.25, ps > 0.2.

Experiment 2 Discussion

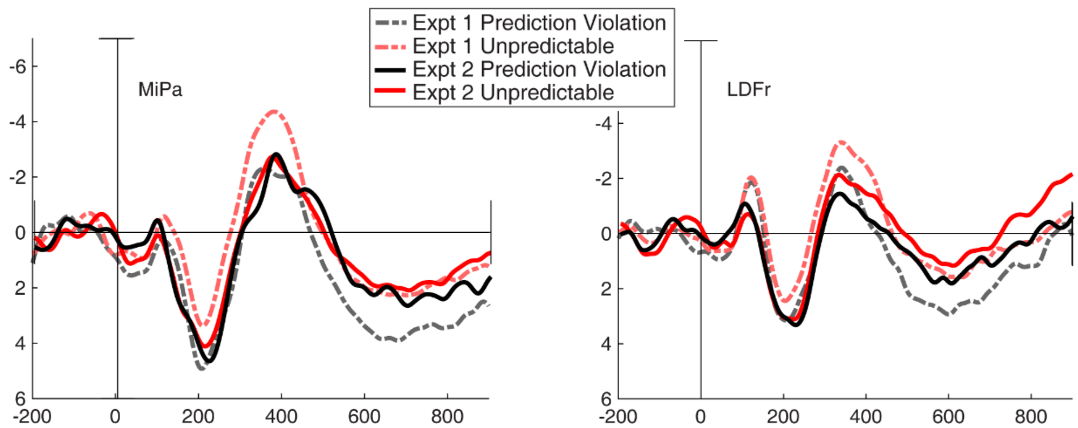
In Experiment 2, participants once again read strongly or weakly constraining sentences ending with an unexpected, but plausible word. The final word would then be seen again three sentences later at the end of a weakly constraining sentence. To encourage the use of prediction during reading, strongly and weakly constraining sentences ending with predicted words from Rommers & Federmeier (2018b) were used as additional fillers. Due to constraints in counterbalancing items across the two stimulus sets (both of which used overlapping materials from Federmeier et al. (2007), the control condition from Experiment 1 was removed, so that the absolute size of the repetition effect could not be ascertained. However, the critical comparison continued to be between the Previously Unpredictable and Previously Violation words, which were the same lexical items repeated in the same weakly constraining contexts.

One goal of Experiment 2 was to determine whether unexpected effect patterns observed for the first presentation sentences in Experiment 1 would replicate in another instantiation of the same paradigm. At the first presentation sentences in Experiment 1, a voltage difference between strongly and weakly constraining sentences was observed beginning at the P2 and continuing through the N400 and late positivity time windows. We did not observe a similar pattern in Experiment 2, beginning with no significant differences observed in the P2 time window, and, in fact, with the numerical pattern reversing (both types of strongly constraining sentences elicited a more positive deflection than weakly constraining sentences). In the N400 time window, ERP patterns at first presentation sentences were aligned with what has previously been seen in the literature. Namely, in comparison to weakly constraining sentences, strongly constraining sentences with unexpected but plausible words showed no difference in N400 amplitude, similar to the results found in Rommers & Federmeier (2018a) and others (e.g., Federmeier et al., 2007; Kutas & Hillyard, 1984). Strongly constraining sentences with predicted endings were also aligned with existing literature in that they showed traditional N400 amplitude reductions (Kutas & Hillyard, 1984; many others) compared to weakly constraining sentences. Finally, in comparison to the final words of weakly constraining sentences, a positive-going deflection over frontal sites in the late time window was seen for unexpected, but plausible words at the end of strongly constraining sentences that was not observed in a similar comparison between weakly constraining sentences and strongly constraining sentences with predicted endings. However, while this effect was not as widely distributed as in Experiment 1, it did still extend beyond frontal sites to include central regions on the left side of the head.

To further compare the constraint-related differences observed in Experiment 1 to results from Experiment 2, we overplotted the waveforms of the two experiments to gain a qualitative sense of how the conditions compare; see Figure 7a. Similar to the comparison between Experiment 1 data and that of Rommers and Federmeier (2018a), this revealed that responses to Unpredictable items in Experiment 1 were more negative on the P2 and N400 and that responses to Prediction Violations in Experiment 1 were more positive over the back of the head in the post-N400 (LPC) time period. Responses over the front of the head, with increased post-N400 frontal positivity to unexpected, but plausible, words in strongly constraining sentences were similar across both experiments and to the pattern in Rommers & Federmeier and in the literature more generally (e.g., Federmeier et al., 2007; Brothers et al., 2015; see also Van Petten & Luka, 2012). The source of the difference in the effect pattern for Experiment 1 compared to both Experiment 2 and Rommers & Federmeier (2018a) is unclear. Experiment 2 contained more prediction confirmations and thus could have better encouraged predictive processing strategies; however, Rommers & Federmeier (2018a) also contained fewer prediction confirmations and showed effects identical to Experiment 2 and different from Experiment 1. Importantly, however, the effect pattern differences at first presentation did not impact the downstream results, as responses for the critical, second presentation words were identical across experiments.

Results from the second presentation sentences were of primary interest for the current study. Comparisons of Previously Violation and Previously Unpredictable items replicated the null effects seen in Experiment 1. There were no significant differences observed on the N400 or the LPC, and results from a paper-and-pencil recognition test once again only showed benefits of repetition and no effect of the context in which the critical word initially appeared. Waveforms from the repetition sentences of Experiment 2 are overplotted with those from Experiment 1, along with the control condition from Experiment 1; Figure 7b. In the N400 time window, waveforms from both conditions in Experiment 2 are largely aligned with their matched counterparts (consisting of the exact same items) from Experiment 1. In comparison to the Not Previously Seen control condition of Experiment 1, the waveforms of both conditions from Experiment 2 have reduced N400 amplitudes, suggesting that the N400 repetition effect seen in Experiment 1 would have held in Experiment 2, which used the exact same items for both the first and second presentations of the critical words. However, because the waveforms from Experiment 2 in the post-N400 time window closely overlap with the waveform of the control condition in Experiment 1, it is not clear whether the LPC repetition effect would have been reobtained. Regardless of this point of uncertainty, what is apparent is the null effect of prior predictability for repetition sentences in both the N400 and the LPC time windows for both experiments, with the N400 repetition patterns in Experiment 2 closely following those observed in Experiment 1. Thus, the critical results from Experiment 1 were therefore reobtained, providing further support for the interpretation that a previous prediction violation has neither a beneficial nor a costly effect on downstream processing.

A



B

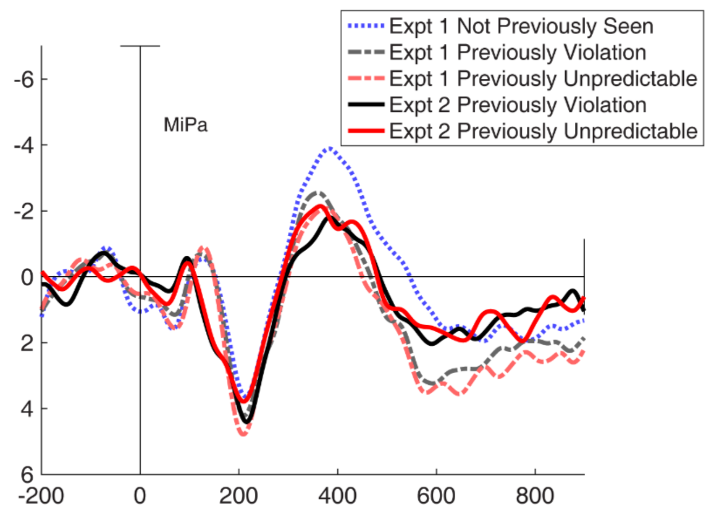


Figure 7. A) Wave forms of the sentence-final words at first presentation from Experiment 1 (dotted-dashed lines) overlain with the first presentations from Experiment 2 (solid lines) at a central site (MiPa) and a frontal site (LDFr). B) Wave forms of the sentence-final words at second presentation from Experiment 1 (dotted and dotted-dashed lines) and Experiment 2 (solid lines) at a central site (MiPa).

Overall Discussion

The goal of this study was to examine the downstream impact of processing a prediction violation, and, more specifically, to determine whether making a prediction and having it violated increased or decreased subsequent memory signals for that violating word. Although historical arguments against prediction have largely been abandoned in the face of both behavioral and electrophysiological evidence demonstrating the facilitative influence of prediction-based expectancy on comprehension, important questions remain about how the language processor copes with the consequences of invalid predictions. These questions are important, especially in light of emerging evidence suggesting that, although prediction plays a role in language processing, it is not ubiquitous. Thus, it becomes important to understand the conditions under which prediction may not be beneficial – such as when it is associated with processing costs. Processing prediction violations is one such condition in which costs of prediction may be incurred, but another possibility is that incorrect predictions are useful for language learning and adaptation, such as in models of implicit learning (e.g., Chang et al., 2000). Thus, an investigation into the mechanism underlying prediction that encapsulates the downstream effects of both valid and invalid predictions is required to understand the role of predictive processing in language comprehension as a whole.

Using the same repetition paradigm as Rommers & Federmeier (2018b, 2018a), in the present pair of experiments, participants read strongly or weakly constraining sentences ending with an unexpected, but plausible, word. Three sentences later the critical word reappeared at the end of a weakly constraining sentence. In Experiment 1, a control condition consisted of the critical word was seen only once, in a weakly constraining sentence. ERP results at the second presentation consistently showed an overall facilitative effect of repetition. However, there was no reliable difference between conditions wherein a word first appeared as a prediction violation or in a non-predictive context on either implicit (N400) or explicit (LPC) measures that have previously been affected by prediction-related processes in prior versions of this paradigm. N400 amplitude was reduced for repeated words, with no significant difference between words that first appeared as prediction violations or as unpredictable words. A similar effect of repetition was found in the increased positivity at the LPC for repeated words, but once again there was no difference between repetition conditions. Results from a paper-and-pencil recognition test conducted after EEG recording followed a similar pattern, with clear benefits of repetition, but no measurable differences between first presentation type.

Because all of the strongly constraining sentences from Experiment 1 ended with unexpected words, prediction might have tended to be discouraged. To investigate the consequences of misprediction under conditions more conducive to predictive processing, Experiment 2 used the same set of sentences as in Experiment 1 but added the strongly constraining sentences ending with predicted words from Rommers & Federmeier (2018b). Again, there were no significant differences found between words that first appeared as prediction violations versus as unpredictable words on either the N400 or the LPC, and results from a paper-and-pencil recognition test again showed benefits of repetition, but no effect from prior predictability. Thus, the lack of effect of prediction violations on downstream implicit and explicit ERP memory effects as well as on behavioral memory performance held across a range of probability of encountering those prediction violations.

There initially appeared to be one difference between the downstream response to previous prediction violations and words that were previously unexpected but not violations of strong expectations. In Experiment 1, a post-hoc analysis of P2 amplitude across frontal and centroparietal channels found a repetition effect that reached significance only for unpredictable words initially appearing in weakly constraining sentences. However, these results were not sustained in Experiment 2, where no significant effects of repetition were found for either of the critical conditions, and numerically the pattern was reversed, with unpredictable words that initially appeared in weakly constraining sentences eliciting a less positive P2 than those that initially appeared as prediction violations.

Overall, these results support the view that encountering a word that violates a prediction built from sentential context ultimately incurs neither cost nor benefit to downstream processing of the unexpected word. Such a finding appears to be in line with the results of Luke & Christianson (2016) and Frisson et al. (2017), who also found no costs of misprediction in late reading measures. Thus, while early skeptics of prediction once argued that language was inherently unpredictable, and that the risk of misprediction could lead to processing costs (e.g., Forster, 1981), this type of experimental evidence suggests that the language processor is able to effectively process unexpected words, even when these constitute prediction violations. Thus, although information about incorrect predictions lingers (as shown by Rommers & Federmeier (2018a)), it does not seem to notably interfere with encoding of the words that replaced those predictions. On the other hand, misprediction does not appear to hold any downstream benefit to memory either, as neither N400/LPC indices nor behavioral memory performance was enhanced for prediction violations.

Results at first presentation in both experiments supported the hypothesis that people were processing the sentences predictively, as there were ERP differences to prediction violations compared with matched unexpected words in weakly constraining sentences that did not permit strong expectations about the sentence-final word. Prior work using these materials has observed a frontal positivity to prediction violations compared to unexpected words in non-constraining contexts (Federmeier et al., 2007), which has since been replicated in other studies as well (DeLong et al., 2014, 2011; Otten & Van Berkum, 2008; Thornhill & Van Petten, 2012). Words in the Prediction Violation condition of Experiment 1 elicited a larger positive-going wave across frontal sites after the N400 than those in the Unpredictable condition; however, a similar effect was also observed over more posterior channels as well. The frontal post-N400 positivity was again found in Experiment 2, this time with less of a distribution across posterior channels. The distribution of the effect suggests that possibility that, in addition to a frontal positivity effect, which has been associated with the processing of *plausible* prediction violations, there was also an LPC-like effect to at least a subset of sentences, perhaps because readers found them relatively implausible. As a whole, however, these differences between the conditions make clear that prediction violations are processed differently than weakly constrained unexpected items when they are first encountered, even though there do not seem to be lasting downstream consequences of those processing differences.

These experiments were designed as follow-up studies to the repetition experiments of Rommers & Federmeier (2018b, 2018a) and, together with that data, present a picture of what the downstream consequences of predicting are like. In Rommers & Federmeier (2018b), strongly and weakly constraining sentences ended with a prediction-confirming word, which then appeared again three sentences later. Results showed a less robust repetition effect at the N400 for repeated words that had initially appeared as prediction confirmations than for those that had appeared in weakly constraining sentences. The repetition effect then disappeared in the LPC, where there was no difference between repeated words that were prediction confirmations and words that were seen for the first time, suggesting that correctly predicting a word led to a more impoverished representation downstream. In another study, Rommers & Federmeier (2018a) changed the endings of their strongly constraining sentences to unexpected but plausible words and left all of their weakly constraining sentences from first and second presentations unchanged. Thus, participants mispredicted words in strongly constraining sentences but then saw their original prediction three sentences later. The results showed what the authors dubbed a "pseudo-repetition" effect: compared to words seen for the first time (and never predicted), N400 amplitude was reduced for words that were being viewed for the first time if they had previously been predicted, although this effect was smaller than that for overt repetition. On the LPC, however, repetition effects only obtained for words that were actually repeated. These findings suggest that violating a prediction was not enough to fully revise expectations and that a representation of the originally predicted word lingered in memory long enough to have an impact three sentences later, on measures linked to implicit – though not explicit – memory. Together the results of these studies suggest that prediction may work to facilitate processing by instantiating upcoming words in ways that have lasting consequences. In the process of such instantiation, the system seems to engage in less thorough processing of an incoming item that matches the prediction. Put another way, while prediction does allow for easier processing, it does so at the expense of the strength of encoding of the incoming word.

Interestingly, whereas the effect of prediction can be seen clearly, both on the processing of that word when obtained and downstream, the effects of prediction on unexpected words may be more short-lived. Although such prediction violations are associated with changes in brain activity when they are encountered, we did not see downstream differences in memory for prediction-disconfirming words compared to words that were simply unexpected. Thus, on the one hand, the lingering activation of the predicted word does not seem to interfere with the representation of the prediction violation that replaced it. On the other hand, the processing reflected in the post-N400-positivities that are elicited by prediction violations – which some have suggested may reflect a conflict detection process and/or a process of revising one's interpretation of the unfolding sentence (Rommers & Federmeier, 2018a, 2018b; Federmeier et al., 2007) – does not seem to augment the encoding of the prediction violation or make it somehow more memorable, at least in the context of ongoing sentence comprehension tasks.

Concluding remarks

In sum, this study extended a recent line of ERP experiments investigating the downstream consequences of valid and invalid predictions. While contextual information has been shown many times to facilitate processing of expected words, it appears to also reduce encoding of the actual input. More specifically, the mechanism of prediction may work via the language processor taking advantage of information from the sentence to ease encoding efforts for a rapid and potentially complex input. A consequence of this reduced encoding can be seen in downstream repetition patterns. Repetition of confirmed predictions leads to less robust repetition effects in the N400 and LPC (Rommers & Federmeier, 2018b), and appearance of a predicted word following a prediction violation leads to a pseudo-repetition effect (Rommers & Federmeier, 2018a). The results of the current study, however, suggest that this impact may be fairly specific to the predicted word, as repetition of a prediction violation yielded implicit, explicit, and behavioral memory effects that were indistinguishable from those to unexpected, non-violations, across two experiments.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, *28*(4), 525–560. https://doi.org/10.1080/01690965.2012.658072

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, *136*, 135–149. https://doi.org/10.1016/j.cognition.2014.10.017

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural Priming as Implicit Learning: A Comparison of Models of Sentence Production. *Journal of Psycholinguistic Research*, *29*(2), 217–230. https://doi.org/10.1023/A:1005101313330

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107. https://doi.org/10.1016/0010-0285(74)90005-X

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, *61*, 150–162. https://doi.org/10.1016/j.neuropsychologia.2014.06.016

DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, *48*(9), 1203–1207. https://doi.org/10.1111/j.1469-8986.2011.01199.x

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 791–801. https://doi.org/10.1037/0278-7393.15.5.791

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*(4), 429–446. https://doi.org/10.1016/0749-596X(88)90066-6

Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*(2), 284–292. https://doi.org/10.1016/j.cognition.2008.11.012

Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H.-J., & Tulving, E. (1997). Event-Related Brain Potential Correlates of Two States of Conscious Awareness in Memory. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(11), 5973–5978.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

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

Federmeier, K. D., & Kutas, M. (1999a). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*(4), 469–495. https://doi.org/10.1006/jmla.1999.2660

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing1. *Cognitive Brain Research*, *8*(3), 373–392. https://doi.org/10.1016/S0926-6410(99)00036-1

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730–747. https://doi.org/10.1016/S0028-3932(01)00193-2

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, *115*(3), 149–161. https://doi.org/10.1016/j.bandl.2010.07.006

Federmeier, K. D., Mai, H., & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition*, *33*(5), 871–886. https://doi.org/10.3758/BF03193082

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133–146.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75–84. https://doi.org/10.1016/j.brainres.2006.06.101

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 1–20. https://doi.org/10.1016/S0022-5371(79)90534-6

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 465–495. https://doi.org/10.1080/14640748108400804

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/a0036756

Frisson, S., Harvey, D. R., & Staub, A. (2017). No prediction error cost in reading: Evidence from eye movements. *Journal of Memory and Language*, *95*, 200–214. https://doi.org/10.1016/j.jml.2017.04.007

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The Search for “Common Sense”: An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience*, *8*(2), 89–106. https://doi.org/10.1162/jocn.1996.8.2.89

Hubbard, R. J., Rommers, J., Jacobs, C. L., & Federmeier, K. D. (in press). Downstream behavioral and electrophysiological consequences of word prediction on recognition memory. *Frontiers in Human Neuroscience*.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, *109*(1), 49–74. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0096-3445.109.1.49

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

Kamide, Y. (2012). Learning individual talkers’ structural preferences. *Cognition*, *124*(1), 66–71. https://doi.org/10.1016/j.cognition.2012.03.001

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55. https://doi.org/10.1023/A:1021933015362

Kazmerski, V. A., & Friedman, D. (1997). Old/new differences in direct and indirect memory tests using pictures and words in within- and cross-form conditions: Event-related potential and behavioral measures. *Cognitive Brain Research*, *5*(4), 255–272. https://doi.org/10.1016/S0926-6410(97)00004-9

Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory & Cognition*, *8*(4), 336–344. https://doi.org/10.3758/BF03198273

Klein, D. E., & Murphy, G. L. (2001). The Representation of Polysemous Words. *Journal of Memory and Language*, *45*(2), 259–282. https://doi.org/10.1006/jmla.2001.2779

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual Distinctiveness Supports Detailed Visual Long-Term Memory for Real-World Objects. *Journal of Experimental Psychology. General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

Kos, M., van den Brink, D., & Hagoort, P. (2012). Individual Variation in the Late Positive Complex to Semantic Anomalies. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00318

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117–129. https://doi.org/10.1016/S0926-6410(03)00086-7

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing. *Electroencephalography and Clinical Neurophysiology. Supplement*, *39*, 325–330.

Kutas, Marta, & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP) [Review-article]. Retrieved May 30, 2017, from Http://dx.doi.org/10.1146/annurev.psych.093008.131123 website: http://www.annualreviews.org/doi/10.1146/annurev.psych.093008.131123

Kutas, Marta, & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. https://doi.org/10.1038/307161a0

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, *61*(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004

Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*, *48*(2), 176–186. https://doi.org/10.1111/j.1469-8986.2010.01058.x

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22–60. https://doi.org/10.1016/j.cogpsych.2016.06.002

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruites on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *59*(6), 477–488. https://doi.org/10.1016/0168-5597(84)90006-6

McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 54–65. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0278-7393.12.1.54

McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science (0956-7976)*, *14*(6), 648–652. https://doi.org/10.1046/j.0956-7976.2003.psci_1480.x

McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*(16), 1735–1751. https://doi.org/10.1016/S0042-6989(03)00237-2

Misra, M., & Holcomb, P. J. (2003). Event–related potential indices of masked repetition priming. *Psychophysiology*, *40*(1), 115–130. https://doi.org/10.1111/1469-8986.00012

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research*, *22*(2), 205–220. https://doi.org/10.1016/j.cogbrainres.2004.08.010

Moss, H. E., & Marslen-Wilson, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1254–1276. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0278-7393.19.6.1254

Ng, S., Payne, B. R., Steen, A. A., Stine-Morrow, E. A. L., & Federmeier, K. D. (2017). Use of Contextual Information and Prediction by Struggling Adult Readers: Evidence From Reading Times and Event-Related Potentials. *Scientific Studies of Reading*, *21*(5), 359–375. https://doi.org/10.1080/10888438.2017.1310213

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/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*(6), 464–496. https://doi.org/10.1080/01638530802356463

Paller, K. A., Kutas, M., & McIsaac, H. K. (1995). Monitoring Conscious Recollection Via the Electrical Activity of the Brain. *Psychological Science (0956-7976)*, *6*(2), 107–111.

Payne, B. R., & Federmeier, K. D. (2017). Pace Yourself: Intraindividual Variability in Context Use Revealed by Self-paced Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, *29*(5), 837–854. https://doi.org/10.1162/jocn_a_01090

Payne, B. R., & Federmeier, K. D. (2018). Contextual constraints on lexico-semantic processing in aging: Evidence from single-word event-related brain potentials. *Brain Research*, *1687*, 117–128. https://doi.org/10.1016/j.brainres.2018.02.021

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, *41*(7), 943–954. https://doi.org/10.1016/S0042-6989(00)00310-2

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191–201. https://doi.org/10.3758/BF03197692

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–509. https://doi.org/10.3758/BF03214555

Rommers, J., & Federmeier, K. D. (2018a). Lingering expectations: A pseudo-repetition effect for words previously expected but not presented. *NeuroImage*, *183*, 263–272. https://doi.org/10.1016/j.neuroimage.2018.08.023

Rommers, J., & Federmeier, K. D. (2018b). Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. *Cortex*, *101*, 16–30. https://doi.org/10.1016/j.cortex.2017.12.018

Rubin, S. R., Petten, C. V., Glisky, E. L., & Newberg, W. M. (1999). Memory Conjunction Errors in Younger and Older Adults: Event-Related Potential and Neuropsychological Data. *Cognitive Neuropsychology*, *16*(3–5), 459–488. https://doi.org/10.1080/026432999380889

Rugg, M. D. (1985). The Effects of Semantic Priming and Word Repetition on Event-Related Potentials. *Psychophysiology*, *22*(6), 642–647.

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.

Schuberth, R. E., Spoehr, K. T., & Lane, D. M. (1981). Effects of stimulus and contextual information on the lexical decision process. *Memory & Cognition*, *9*(1), 68–77. https://doi.org/10.3758/BF03196952

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 344–354. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0278-7393.14.2.344

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, *24*(2), 232–252. https://doi.org/10.1016/0749-596X(85)90026-9

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 658–672. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0096-1523.7.3.658

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, *112*(1), 1–36. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0096-3445.112.1.1

Sussman, R. S., & Sedivy, J. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, *18*(2), 143–163. https://doi.org/10.1080/01690960143000498

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Taylor, W. L. (1953). "Cloze Procedure": A New Tool for Measuring Readability. *Journalism Bulletin*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, *83*(3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007

Traxler, M. J., & Foss, D. J. (2000). Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1266–1282. http://dx.doi.org.proxy2.library.illinois.edu/10.1037/0278-7393.26.5.1266

Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word recognition processes. *Perception & Psychophysics*, 409–420.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating Upcoming Words in Discourse: Evidence From ERPs and Reading Times. *Journal of Experimental Psychology. Learning, Memory & Cognition*, *31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Van Petten, C., & Senkfor, A. J. (1996). Memory for words and novel visual patterns: Repetition, recognition, and encoding effects in the event-related brain potential. *Psychophysiology*, *33*(5), 491–506. https://doi.org/10.1111/1469-8986.ep10969431

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the Word Repetition Effect with Event-Related Potentials. *Journal of Cognitive Neuroscience*, *3*(2), 131–150. https://doi.org/10.1162/jocn.1991.3.2.131

West, R. F., & Stanovich, K. E. (1978). Automatic Contextual Facilitation in Readers of Three Ages. *Child Development*, *49*(3), 717–727. https://doi.org/10.1111/1467-8624.ep10425716

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Wlotko, E. W., & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, *45*(13), 3001–3014. https://doi.org/10.1016/j.neuropsychologia.2007.05.013

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, *62*(1), 356–366. https://doi.org/10.1016/j.neuroimage.2012.04.054

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975–988. https://doi.org/10.1037/a0029206

Wlotko, E. W., Lee, C.-L., & Federmeier, K. D. (2010). Language of the aging brain: Event-related potential studies of comprehension in older adults. *Language and Linguistics Compass*, *4*(8), 623–638. https://doi.org/10.1111/j.1749-818X.2010.00224.x

Woollams, A. M., Taylor, J. R., Karayanidis, F., & Henson, R. N. (2008). Event-related Potentials Associated with Masked Priming of Test Cues Reveal Multiple Potential Contributions to Recognition Memory. *Journal of Cognitive Neuroscience*, *20*(6), 1114–1129. https://doi.org/10.1162/jocn.2008.20076