Predictability and novelty in literal language comprehension: An ERP study

Tristan Davenport\textsuperscript{a,*}, Seana Coulson\textsuperscript{a,b}

\textsuperscript{a}Cognitive Science Department, University of California, San Diego, United States
\textsuperscript{b}Center for Research in Language, University of California, San Diego, United States

ARTICLE INFO

Article history:
Accepted 15 July 2011

Keywords:
Conceptual integration
LPC
N400
Semantics
P600
Pragmatics

ABSTRACT

Linguists have suggested that one mechanism for the creative extension of meaning in language involves mapping, or constructing correspondences between conceptual domains. For example, the sentence, "The clever boys used a cardboard box as a boat," sets up a novel mapping between the concepts cardboard box and boat, while "His main method of transportation is a boat," relies on a more conventional mapping between method of transportation and boat. To examine the electrophysiological signature of this mapping process, electroencephalogram (EEG) was recorded from the scalp as healthy adults read three sorts of sentences: low-cloze (unpredictable) conventional ("His main method of transportation is a boat,"), low-cloze novel mapp'ing ("The clever boys used a cardboard box as a boat,"), and high-cloze (predictable) conventional ("The only way to get around Venice is to navigate the canals in a boat.").

Event-related brain potentials (ERPs) were time-locked to sentence final words. The novel and conventional conditions were matched for cloze probability (a measure of predictability based on the sentence context), lexical association between the sentence frame and the final word (using latent semantic analysis), and other factors known to influence ERPs to language stimuli. The high-cloze conventional control condition was included to compare the effects of mapping conventionality to those of predictability.

The N400 component of the ERPs was affected by predictability but not by conventionality. By contrast, a late positivity was affected both by the predictability of sentence final words, being larger for words in low-cloze contexts that made target words difficult to predict, and by novelty, as words in the novel condition elicited a larger positivity 700–900 ms than the same words in the (cloze-matched) conventional condition.

1. Introduction

The meaning activated in response to a word is heavily influenced by the context in which that word is embedded. For example, boat depicts a very different sort of object in the context of, say, a cruise ship vacation, and that of the bathtub. Linguists have suggested that one mechanism for the creative extension of meaning in language involves mapping, or constructing correspondences between conceptual domains. Lee (2002), for example, notes that the expression rumor mill involves a mapping between the domain of mills and that of social talk that includes correspondences between the mill (such as a flour mill) and the people who gossip, the process of industrial production and the talk, as well as between the product (e.g. flour) and the rumor. While Lee’s example is clearly figurative, mappings have also been shown to form an
important component of understanding fully literal language (Coulson and Oakley, 2000; Fauconnier, 1997). For example, the sentence, “The clever boys used a cardboard box as a boat,” sets up a novel mapping between the concepts cardboard box and boat. In this context, boat must be interpreted as bearing some of the semantic features of cardboard boxes and lacking some of the features most commonly associated with boats.

The neurocognitive underpinnings of such meaning accommodation are largely unknown, as research on this topic has focused primarily on how the predictability of a word in its context affects the ease of semantic processing, as indexed by the amplitude of the N400 component in the event-related brain potential (ERP) (Kutas and Federmeier, 2000; Kutas and Hillyard, 1980, 1984). A negative-going peak in ERPs elicited by words and other sorts of meaningful stimuli, the N400 is interpreted as indexing either the activation (Lau et al., 2009) or integration (Chwilla and Kolk, 2005) of word meaning. N400 amplitude is systematically reduced by facilitative sentence context, and is inversely correlated with cloze probability, a measure of the predictability of a given word in a particular sentence context (DeLong et al., 2005). Beyond predictability, N400 amplitude is also influenced by the nature of a word’s relationship to its context, including lexical association (Kutas and Hillyard, 1983), taxonomic category membership (Federmeier and Kutas, 1999a, 1999b), and contextually evoked frames, scripts, or schemas (Bicknell et al., 2010; Chwilla and Kolk, 2005; Coulson and Kutas, 2001).

A recent ERP study of word processing in context examined plausible and implausible novel language as a test of embodied theories of processing (Chwilla et al., 2007). Participants read short vignettes in which the final sentence described a novel way of performing a familiar action (… They let the canoe into the water and paddled with frisbees/ pullovers). The critical word rendered the action either plausible (frisbees) or implausible (pullovers). The larger N400 elicited by the implausible words was taken as evidence in favor of an embodied account of meaning: comprehenders activated motor and perceptual information in the course of processing, and this activation was more effortful when the physical situation described was impossible to effect in real life.

1.1. Figurative language and semantic mapping

A type of semantic mapping that has been relatively well explored with the ERP method is metaphor, with psycholinguistic research focusing particularly on the contrast between metaphorical and literal language. An oft-cited traditional view of metaphoric meaning construction is that the cognitive system first tries to construct a literal interpretation of the word or sentence, and then, if that fails, a metaphorical meaning is constructed instead (Grice, 1975). This hypothesis was found wanting when Gibbs (1994) found no significant reading time difference between literal sentences and sentences containing conventional metaphors. An ERP investigation (Pynte et al., 1996) supported Gibbs’s (1994) hypothesis that when context supports a metaphorical interpretation, a metaphorical mapping between domains can be accessed without the cognitive system having to first consider an anomalous literal interpretation of the word.

Coulson and Van Petten (2002) conducted an ERP study comparing novel metaphors, conventional literal language, and novel literal language (i.e. “literal mapping”), in order to study the neural correlates of constructing different kinds of semantic mappings. The novel literal stimuli were created by using the same critical words as in the other conditions, but in literal contexts that enforced novel interpretations of them, such that they bore some semantic features of a concept referenced earlier in the sentence. For instance, the word boat in the novel literal stimulus, “The clever boys used a cardboard box as a boat,” must be interpreted as referring to an object bearing some of the semantic features of boats and some of the semantic features of boxes. By contrast, “His main method of transportation is a boat” enforces a more conventional interpretation of the critical word.

Coulson and Van Petten (2002) observed a graded N400 effect, smallest for conventional literal, marginally larger for novel literal, and largest for novel metaphors. They also observed enhanced late positivities of distinct scalp distributions to both the novel literal and novel metaphorical conditions. The novel literal condition elicited a frontal late positivity, while the novel metaphor condition elicited a posterior effect of similar size and latency. Because all of the conditions were matched for cloze probability, and because all of the critical words were repeated and counterbalanced across conditions, it must have been the case that some other aspects of the critical word’s relationship to its context that modulated the difficulty of processing, as indexed by the N400 and late positivities.

Coulson and Van Petten’s (2002) analysis of the late positivity elicited by low-frequency words appearing in highly constrained contexts (Van Petten, 1993), interpreted as reflecting more intense processes of semantic memory access related to the more precise and/or complex meanings of low-frequency words (Zipf, 1945). Coulson and Van Petten’s (2002) analysis of the late positivity elicited by metaphors is essentially that a similar frequency effect exists for concepts or semantic features. The late positivity elicited by a word in a novel metaphorical context reflects the activation of semantic features which are not typically associated with that word-form and which are necessary to complete the conceptual blend (Fauconnier and Turner, 2002) and interpret the metaphor.

Interpretation of Coulson and Van Petten’s (2002) findings thus depends on establishing the functional significance of the late positivities elicited by words with novel literal versus metaphorical meanings, and their relationship to ERP effects typically elicited for semantic anomalies. Besides the N400, anomalous sentence completions sometimes elicit a late positive-going response (e.g. Ford et al., 1996; Gunter et al., 1992; Schwartz et al., 1996; Woodward et al., 1993), referred to as the P600 (Kuperberg, 2007), the late positive complex (LPC, Federmeier et al., 2007), and the post-N400 positivity (Van Petten and Luka, 2006). The so-called “semantic P600” (Kuperberg, 2007) is largest posteriorly, and has a topography similar to the ERP effect that was traditionally associated with the detection of dispreferred or illegal syntactic structures (Coulson et al., 1998; Kuperberg et al., 2006; Osterhout et al., 1994). Another late
positivity shows a consistent frontal distribution (the “frontal positivity”), and is elicited by unexpected words in highly predictable contexts. It is therefore thought to index the neural costs associated with suppressing or revising a failed prediction (Federmeier et al., 2007).

Prior studies have compared the neural costs of conventional and novel semantic mappings in comprehending literal language (Coulson and Van Petten, 2002) as well as metaphorical language (Arzouan et al., 2007; Lai et al., 2009). These and similar studies have generally observed that semantically novel stimuli elicit larger N400 responses than semantically conventional stimuli, as well as in some cases larger late positivities. Both the N400 and a frontal late positive component have been observed in studies that manipulate the predictability of critical words (DeLong et al., in press; Federmeier et al., 2007; Kutas and Hillyard, 1984). However, no study as yet has directly compared the ERP effects of a novel/conventional manipulation with those of a predictability manipulation. The present study addresses that gap in the literature by investigating the neural responses to both of these variables.

1.2. The present study

The present study seeks to replicate and extend the work of Coulson and Van Petten (2002), comparing the ERP effects of predictability and semantic conventionality differences, on the brain response to novel versus conventional literal language. To that end, we recorded ERPs as participants read materials from Coulson and Van Petten’s (2002) conventional and novel literal conditions, as well as a high-cloze conventional condition (see Table 1 for sample materials and Fig. 1 for a summary of the experimental paradigm). Observed differences in N400 amplitude in the earlier study, for example, might be attributable to subtle differences in the semantic distance between the critical words and the preceding sentence context in the (conventional) literal versus the (novel) literal mapping conditions (as noted in Coulson and Matlock (2001)). Consequently, the present study employed a subset of the materials for which this factor was closely matched. Amplitude differences in the N400 elicited by conventional and novel mappings would thus suggest that mapping complexity affects the meaning activation and integration processes indexed by that component. Moreover, the design of the present study allowed us to compare the size and topography of late positivities elicited by novel literal words to those elicited by conventional uses of the same words.

Another goal of the present study is to assess Coulson and Van Petten’s (2002, footnote 6) speculation that the frontal positivity elicited by novel literal stimuli reflected the large number of mental state predicates in that condition, and that interpreting the novel literal sentences may have involved considerably more theory of mind processing than the other two conditions. To test this hypothesis, we divided our novel mapping stimuli into categories based on the relationship between the target word and the previously mentioned concept onto which its meaning is mapped. As outlined in greater detail in Section 4, materials were divided into “Mental” sentences in which the semantic mapping relied on the juxtaposition of two mental states, and “Physical” sentences in which the semantic mapping relied on differences in the surface or functional features of objects. These Mental and Physical categories were carefully balanced for factors known to influence the brain response to words, including cloze probability and semantic distance from the sentence context as measured by latent semantic analysis (LSA) (Landauer and Dumais, 1997; Landauer et al., 1998). If the anterior distribution of the late positivity reflects neural generators dedicated to theory of mind processing, we might expect Mental mappings to elicited a late positivity largest over anterior regions and Physical mappings to elicited a more posterior positivity, or no enhanced positivity at all.

Overall, differences in ERPs elicited by critical words in novel literal contexts relative to the same words in cloze-matched conventional contexts would indicate the brain’s sensitivity to our manipulation of mapping complexity. One tangible manifestation of this difference involves semantic features of the critical word activated in the conventional compared to the novel mapping condition. In a study in which participants were asked to list the features of the critical word presented alone and in our sentence contexts, Coulson and Matlock (2001) found that the features listed for words in the null context were more similar to those listed for words in the conventional than in the novel literal contexts. Any observed ERP effects of novelty presumably index processes related to semantic memory access, during which time the cognitive system activates semantic features not typically activated in response to that particular word.

2. Results

2.1. Comprehension

Following each experimental trial were two checks for comprehension. First, participants were cued by the word READ? to indicate by button-press whether they had read the preceding stimulus sentence. Participants indicated on 94% of trials that they had read the stimulus sentence. An analysis of variance (ANOVA) with sentence type as a factor revealed no significant difference among stimulus types, [F(2,51) = 0.75, p = 0.48].

Table 1 – Sample stimuli.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High-cloze conventional (HC)</strong></td>
<td>To buy bread you go to baker, and to buy meat you go to a butcher.</td>
</tr>
<tr>
<td>T/F: Butchers and bakers sell different products.</td>
<td>Experts say the main source of taxes in this town is the factory.</td>
</tr>
<tr>
<td>T/F: Crops and machines come from the same place.</td>
<td>Crops come from a farm, but manufactured goods like cars come from a factory.</td>
</tr>
<tr>
<td><strong>Low-cloze conventional (LC)</strong></td>
<td>Down the street from the vegetable stand is a butcher.</td>
</tr>
<tr>
<td>T/F: Vegetarians and meat-eaters can shop on the same street.</td>
<td>The factory operates tax-free.</td>
</tr>
<tr>
<td><strong>Low-cloze novel (LN)</strong></td>
<td>During the war, the Jewish surgeon had to work as a butcher.</td>
</tr>
<tr>
<td>T/F: He couldn’t operate on patients during the war.</td>
<td>He had apparently seen the artists’ studio but thought it was a factory.</td>
</tr>
<tr>
<td>T/F: The studio looked very industrial from the outside.</td>
<td>Low-cloze novel (LN)</td>
</tr>
</tbody>
</table>
Participants were also presented with a true/false comprehension question after each trial. Correct responses to the comprehension questions indicated that participants had understood the preceding stimulus. Participants correctly answered 92% of comprehension questions, with no effect of sentence type, \(F(2,51)=1.8, p=0.19\).

2.2. Event-related potentials

Fig. 2 shows the comparison between the high-cloze (HC) and each of the different low-cloze (LC and LN) sentence conditions, and Fig. 3 shows the direct comparison between conventional and novel language in unpredictable contexts (LC vs. LN).

2.2.1. Sentence Type analysis

Sentence Type analyses involved measurements of ERP components known to be influenced by the predictability of sentence final words, including the P2, N400, as well as the initial (500–700 ms) and later (700–900 ms) phases of the post-N400 positivity. Measurements were subjected to repeated measures ANOVA with factors Sentence Type (High Cloze, Low Cloze Conventional, Low Cloze Novel) and Electrode (29 sites). Results of these analyses are presented in Table 2. Effects of Sentence Type were observed on the N400, and both measurements of the late positivity. As can be seen in Fig. 2, High Cloze sentences elicited a less negative N400 than did either of the Low Cloze sentences (LC or LN). As is typical for the N400 (see Kutas and Federmeier (2011) for review), the N400 effect was broadly distributed, and largest over centro-parietal electrode sites (see Fig. 2).

Between 500 and 700 ms post-stimulus onset, both low-cloze conditions (LC and LN) elicited ERPs more positive than did the high cloze condition. Main effects of sentence type were significant for both the HC vs. LC comparison \(F[1,17]=5.5, p=0.03; \text{effect size}=1.1 \mu V\) and for the HC vs. LN comparison \(F[1,17]=5.7, p=0.03; \text{effect size}=1.1 \mu V\). Follow up tests were conducted on medial electrodes (F3/4, FC3/4, C3/4, CP3/4, P3/4, O1/2) in order to test for interactions between sentence type, anteriority, and hemisphere. In the HC vs. LC comparison, a significant Sentence Type × Anteriority effect reflected the frontal focus of this positivity \(F[5,85]=11.9, p=0.0001\). The HC vs. LN comparison revealed significant interactions between Sentence Type and Anteriority \(F[5,85]=6.3, p<0.01\), Sentence Type and Hemisphere \(F[1,17]=16.7, p<0.001\) and a three-way interaction between Sentence Type,
Anteriority, and Hemisphere \((F[5,85]=3.1, p=0.04)\), reflecting its right anterior focus.

Between 700 and 900 ms post-stimulus onset, both low-cloze conditions (LC and LN) elicited more positive ERPs than the HC condition. Main effects of sentence type were significant for the HC vs. LC comparison \((F[1,17]=12.4, p<0.005; \text{effect size}=1.6 \mu V)\) and for the HC vs. LN comparison \((F[1,17]=21.2, p<0.0005; \text{effect size}=2.5 \mu V)\). Pairwise comparisons of the HC and LC conditions on medial sites failed to reveal an Anteriority×Sentence Type interaction \((F[5,85]=1.6, p=0.23)\). However, a similar analysis of the HC and LN conditions on medial sites revealed a robust Anteriority×Sentence Type interaction \((F[5,85]=4.9, p=0.02)\), reflecting the anterior focus of this effect (see Fig. 2).

2.2.2. Conventionality: low cloze conventional versus low cloze novel

Besides the Sentence Type analysis described above, additional analyses were conducted to compare the amplitude and scalp distribution of ERPs elicited by the two low-cloze conditions (LC and LN). Results of these planned comparisons are presented in Table 3. In contrast to the analysis of predictability (Section 2.2.1), no effects were observed on either the N400 or the initial phase of the post-N400 positivity.

Fig. 2 – Cloze comparison. The left-hand column shows ERPs from three midline scalp channels time-locked to the final words in high cloze conventional (solid black) and low cloze conventional (dotted blue) sentences. Negative voltage is plotted upwards. Adjacent scalp maps show the topography of the difference wave formed by subtracting ERPs to high cloze conventional from low cloze conventional at 400 ms post-stimulus (top), 600 ms post-stimulus (middle), and 800 ms post-stimulus (bottom). Analogously, the right-hand column shows ERPs from the same three midline scalp channels time-locked to final words in high cloze conventional (solid black) and low cloze novel (dashed blue) sentences, alongside isovoltage maps of their difference waves at 400, 600, and 800 ms post-stimulus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
By contrast, measurements of ERPs elicited 700–900 ms post-onset revealed a robust effect of Conventionality, as low cloze novel stimuli elicited more positive ERPs than did their low cloze conventional counterparts (see Table 3; Fig. 3). Although the overall analysis did not indicate significant differences in the scalp distribution of the positivities elicited by the LC and LN conditions, inspection of Fig. 4 suggested the conventionality effect had a right anterior focus. The topographic profile of this effect was further pursued via separate post hoc analyses of left and right medial sites. Analysis of the right medial electrodes (F4, FC4, C4, CP4, P4, O2) revealed a marginal main effect of Conventionality (F[1,17], p=0.057), and a significant Conventionality × Anteriority interaction (F[5,85]=5.2, p=0.01), reflecting its restricted occurrence over frontal sites (see Fig. 3). A similar analysis of left medial sites yielded only a main effect of Conventionality (F[1,17]=4.2, p<0.01), but no interaction between Conventionality and Anteriority (F[5,85]<1). Results of these post hoc topography analyses must, however, be regarded with caution.

2.2.3. Semantic mapping types
Planned comparison contrasted mean amplitude measurements of ERPs elicited by LN stimuli in the Mental and Physical categories versus their LC counterparts. Measurement intervals for the N400 and post-N400 positivity were as noted above, (300–500 ms, 500–700 ms, and 700–900 ms after stimulus onset). Analysis included factors Mapping (Mental vs. Physical), Sentence (LN/LC), and electrode site (29 levels). There were no main (F[1,17]=2.1, p=0.16) or interaction effects of semantic mapping types (Mental vs. Physical) on the N400 (Mapping×Sentence: F[1,17]<1; Mapping×Sentence×Electrode: F[28,476]=1.34, p=0.26). There were also no significant main or interaction effects of mapping type (Mental vs. Physical) on either the early portion of the late positivity (Mapping: F[1,17]<1; Mapping×Sentence: F[1,17]<1; Mapping×Sentence×Electrode: F[28,476]=1.4, p=0.55), or on the later part of the late positivity (Mapping: F[1,17]=1.16, p=0.3; Mapping×Sentence: F[1,17]<1; Mapping×Sentence×Electrode: F[28,476]<1).

3. Discussion
The present study compared ERP effects of predictability and semantic conventionality for sentence-final words in three kinds of literal sentences: a high-cloze (predictable)
conventional (HC) mapping condition such as "A giant, ornate church like Westminster Abbey is called a cathedral," a low-cloze (unpredictable) conventional (LC) mapping condition such as "It scared him to be alone in the cathedral," and a low-cloze (unpredictable) novel (LN) mapping condition, such as, "At one time, this movie house was a cathedral." The latter condition involved contexts that require language users to understand a mapping, or correspondence, between the target word and a concept evoked by a prior word in the sentence. For example, the transformation context in "At one time, this movie house was a cathedral," promotes a mapping between the concept movie house (the current function of the building) and cathedral (the prior function of the building). This mapping in turn can promote pragmatic inferences about the features of the movie house, such as the idea that it is a large building with high ceilings. The primary aim of the study was thus to compare the ERP effects of predictability and semantic mapping conventionality (discussed in Sections 3.1 and 3.2 below).

A secondary aim was to test whether the type of novel semantic mapping (namely, whether it involved a mapping between mental versus physical states) affected the topography of the late positivities elicited by the LN condition. However, as noted in Section 2.2.3, comparison of LN Mental stimuli to their LC counterparts revealed a late positivity of similar size and distribution to our comparison of the LN Physical stimuli to their LC counterparts. Results are thus inconsistent with Coulson and Van Petten's (2002) suggestion that the right anterior distribution of the late positivity associated with LN (literal mappings) was due to the use of brain regions implicated in Theory of Mind. Discrepancies between the present study and the earlier one by Coulson and Van Petten (2002) might be due to differences in the surrounding materials, such as the inclusion here of high cloze (predictable) sentences, or the inclusion of metaphorical sentences in the prior study. Differences in the global characteristics of the two studies might have modified participants’ interpretive strategies in a way that altered the post-N400 positivities observed. We find it more likely, however, that Coulson and Van Petten’s (2002) post hoc comparison of sentences which either did or did not require theory-of-mind had subtle linguistic confounds not present in the more carefully controlled comparison here between Mental versus Physical mappings.

### Table 2 – Effects of Sentence Type (all 3 conditions).

<table>
<thead>
<tr>
<th>Time window</th>
<th>Factors</th>
<th>F</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2 (200–300)</td>
<td>Sentence</td>
<td>F(2,34)=1.1</td>
<td>0.31</td>
</tr>
<tr>
<td>All 3 conditions</td>
<td>Sentence</td>
<td>F(1,17)=3.1</td>
<td>0.097</td>
</tr>
<tr>
<td>HC vs. LC</td>
<td>Sentence</td>
<td>F(1,17)=0.01</td>
<td>0.94</td>
</tr>
<tr>
<td>N400 (500–500 ms)</td>
<td>Sentence</td>
<td>F(2,34)=4.2</td>
<td>0.051</td>
</tr>
<tr>
<td>All 3 conditions</td>
<td>Sentence × electrodes</td>
<td>F(56,952)=7.4</td>
<td>0.0001***</td>
</tr>
<tr>
<td>HC vs. LC</td>
<td>Sentence</td>
<td>F(1,17)=7.4</td>
<td>0.01*</td>
</tr>
<tr>
<td>HC vs. LN</td>
<td>Sentence</td>
<td>F(1,17)=11.5</td>
<td>&lt;0.0001***</td>
</tr>
<tr>
<td>Sentence × electrodes</td>
<td>F(28,476)=7.8</td>
<td>0.002**</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3 – Effects of novelty/conventionality (LC vs. LN).

<table>
<thead>
<tr>
<th>Time window</th>
<th>Factors</th>
<th>F</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2 (110–250 ms)</td>
<td>novel/conventional</td>
<td>F(1,17)=1.7</td>
<td>0.21</td>
</tr>
<tr>
<td>N400 (500–500 ms)</td>
<td>novel/conventional</td>
<td>F(1,17)=2.6</td>
<td>0.12</td>
</tr>
<tr>
<td>Early LPC (500–700 ms)</td>
<td>novel/conventional</td>
<td>F(1,17)=0.0</td>
<td>0.96</td>
</tr>
<tr>
<td>Later LPC (700–900 ms)</td>
<td>novel/conventional</td>
<td>F(1,17)=7.5</td>
<td>0.01*</td>
</tr>
</tbody>
</table>

*=p<0.05; **=p<0.01; ***=p<0.001.

![Fig. 4](image_url)
the target word, resulted in modulation of the N400 component, in keeping with a large literature on sentence context effects (reviewed in Kutas and Federmeier, 2011). Although the high-cloze condition differed from the low-cloze conditions in contextual constraint, a variable thought to affect the P2 component (Wlotko and Federmeier, 2007), no P2 effect of sentence type was observed. Sentence final words in the high-cloze conventional (HC) condition elicited a smaller N400 than did either of the low-cloze conditions, as well as a less positive response in the 500–900 ms window (see Fig. 2). The latter effect resembles the late positivity observed by Federmeier et al. (2007) to unpredictable words embedded in highly predictive contexts, as well as the late positivity reported by DeLong et al. (in press) to more versus less expected (plausible) sentence completions. Although the difference in mean cloze probabilities between the HC condition (mean cloze = 0.67) and the low-cloze conditions (mean cloze = 0.08 and 0.09) is not as great as that used by Federmeier et al. (2007), it is similar to the range of cloze differences reported by DeLong et al. (in press).

The N400 indexes the difficulty of activating conceptual knowledge associated with a word, which is affected both by properties of the word itself, and the context in which it appears (Van Petten and Luka, 2006). As noted in the introduction, the functional significance of the late positive-going response observed is not well established, though DeLong et al. (in press) suggest a relationship between its amplitude and the need to suppress or revise erroneous predictions about upcoming words.

Though both the N400 and the late positivity are elicited by manipulations of the predictability of words in context, differences in the polarity and latency of these components suggest that different brain regions underlie their generation. Moreover, in an ERP study comparing sentence context effects in patients with frontal lesions and age-matched controls, Swick (2004) and Swick et al. (1998) found very similar N400 effects in patients and controls followed by a late positivity which was present only in controls. Its absence in the patient group suggests the late positivity has a generator in the frontal electrodes that contrasts with the much larger positive-going effect over the rest of the scalp. Moreover, the same anterior scalp regions that show a small positive response during the N400 time window, are also the focus of the much larger and more broadly distributed late positivity at 800 ms (see Fig. 2), as well as the late positive effect of mapping conventionality (see Fig. 4). Overlapping activity between the generators of the N400 and those underlying the positive-going effect would serve to attenuate the amplitude of the N400, thereby obscuring small predicted amplitude differences between the two low-cloze conditions (LC and LN).

3.2. Semantic conventionality

Comparison of the two low-cloze conditions, low-cloze conventional (LC) and low-cloze novel (LN), revealed amplitude differences only in the latter part of the late positivity (700–900 ms after the onset of the target words). The absence of a significant N400 effect differs somewhat from findings reported by Coulson and Van Petten (2002) in a similar study. They found that LN stimuli (referred to by the authors as the “literal mapping” condition) elicited numerically larger N400 components than did LC stimuli (referred to as the “literal” condition). However, pairwise comparison of N400 amplitudes in the literal versus literal mapping conditions failed to reach significance, as did comparison of their literal mapping and metaphor conditions (Coulson and Van Petten, 2002).

The similar-sized N400 elicited by LC and LN conditions in the present study may be due, at least in part, to the concurrent presence of a positive-going response in the ERPs beginning sometime 300–500 ms after stimulus onset. Inspection of the two 400 ms post-onset scalp maps in Fig. 2 both suggest a small positive-going predictability difference over frontal electrodes that contrasts with the much larger negative-going effect over the rest of the scalp. Moreover, the same anterior scalp regions that show a small positive response during the N400 time window, are also the focus of the much larger and more broadly distributed late positivity at 800 ms (see Fig. 2), as well as the late positive effect of mapping conventionality (see Fig. 4). Overlapping activity between the generators of the N400 and those underlying the positive-going effect would serve to attenuate the amplitude of the N400, thereby obscuring small predicted amplitude differences between the two low-cloze conditions (LC and LN).

The larger late positivity observed here for the LN than for the LC condition replicates that reported by Coulson and Van Petten (2002), who used the same low-cloze novel condition. They did not analyze it beyond noting its presence, however, reserving most of their discussion for a posterior positivity elicited by novel metaphorical stimuli. The right anterior focus of the conventionality effect observed in the present study (see Fig. 4) resembles that in Coulson and Van Petten (2002), as well as the late predictability effect (500–900 ms) in the present study (see Fig. 2). The novelty/conventionality effect observed in the present study is also similar to the right-central late positivity.
elicited by sentence-final words in metaphorical compared to literal contexts (e.g., *Unemployment/Cholera is a plague*, De Grauwe et al., 2010). Results of our statistical analysis suggest that the LN materials elicited a slightly larger positivity than did the LC, but with a similar scalp distribution. As such, it can be understood as a modulation of the positive-going predictability effect described in Section 3.1.

De Grauwe et al. (2010) propose that the enhanced late positivity to metaphors indexes the conflict between the implausibility of the sentence’s literal meaning and its metaphorical meaning. This analysis implies that embedding unusual semantic mappings in literally plausible sentences would remove the conflict and consequently remove the late positivity. However, the present results are inconsistent with that prediction. Here we employed novel semantic mappings embedded in literally plausible sentences and observed a similarly-distributed late positivity that could not be explained by differences in cloze probability or LSA relatedness.

### 3.3 Late positivities

We suggest that a better explanation of the late positivities observed in studies of novel semantic mappings, both literal and metaphorical, is an account based on expectancy for higher-order semantic relationships. We begin with the assumption that a comprehender incrementally constructs a situation model informed by world knowledge to represent a given sentence’s meaning (Zwaan, 2009). The evolving situation model creates a set of linguistic expectations, facilitating the processing of words that fit the situation model, and causing conflicts for words that do not. In these cases, the semantic knowledge held in working memory must often be suppressed or revised, a process indexed by the late positivity.

Similar late positivities were presumably observed in the present study and that by De Grauwe et al. (2010) because sentence final words involving both novel literal and metaphorical words are likely to require revision of the reader’s situation model and consequent linguistic expectations. For example, “Unemployment is a...” creates an expectation for words conventionally used to describe unemployment, which must be revised when “plague” is encountered (from De Grauwe et al., 2010). Similarly, “The clever boys used a cardboard box as a...” creates an expectation for words describing the conventional uses of a cardboard box. This expectation must also be revised when “boat” is encountered.

In both cases, the revision of the expectation is reflected in a late positivity. The positivity derives not from the word’s figurativity, but rather from its degree of conflict with higher-order expectancies. Accordingly, metaphorical words that are predictable in context, as in conventional metaphors and idioms, have not been shown to elicit late positivities (Arzouan et al., 2007; Lai et al., 2009).

Violations of high-level semantic and pragmatic expectations has often elicited late positive effects in ERP language studies. For example, late positivities of varying topographies have been reported in ERP studies of jokes that require reinterpretation of the extant situation model (Coulson and Kutas, 2001; Coulson and Lovett, 2004; Coulson and Williams, 2005; Coulson and Wu, 2005). For example, Coulson and Kutas (2001) observed larger late positivities to joke (wife) than non-joke (daiquiri) endings of sentences such as (e.g., *When I asked the bartender for something cold and full of rum, he recommended his...*), even though sentence-final words were equally unpredictable. Moreover, ERP studies of irony report that the “trigger” of an ironic utterance elicits a larger late positivity than non-ironic control stimuli (Cornejo et al., 2007; Regel et al., 2009, 2010).

Consistent with our hypothesis that these post-N400 positivities reflect the violation of expectancies, experimental manipulations that affect listeners’ high level expectancies have been shown to modulate their amplitude. For example, the irony-related P600 can be modulated by the type of comprehension task (Cornejo et al., 2007), the overall frequency of ironic sentences (Regel et al., 2009), and the frequency with which a particular speaker uses ironic language (Regel et al., 2010). Although a parallel streams analysis (Kuperberg, 2007) could explain the general finding, it could not explain the susceptibility of the irony P600 to manipulations of frequency and predictability.

A more nuanced account of these data might be found in an account based on expectations in the form of an evolving situation model. The critical word in a sentence such as, *This is so informative*, (from Regel et al., 2010) could be easy or difficult to integrate with the situation model in its current state, depending on whether the discourse context describes an informative lecture or a vacuous one. However, in both cases, the subject might be expecting an adjective related to the quality of the lecture, and this thematically relevant word could account for the lack of an N400 effect on the ironic word.

Linguistic expectancies, of course, derive from multiple sources, and are by no means limited to the situation model. Expectations might arise from associative semantic relationships, from structures of world knowledge (Elman, 2009), from pragmatic expectations about speakers’ communicative intentions (Regel et al., 2010), from implicit knowledge of the conditional probabilities of words and phrases, possibly in the form of constructions (e.g., Goldberg, 1995), and even from expectations that discourse will continue in the same language instead of switching (see Moreno et al., 2002, for posterior and frontal late positivities elicited by code switching in idiom contexts). Comprehenders may form narrow expectations of single words in high-constraint contexts (Federmeier et al., 2007), as well as graded semantic expectations that may be violated to different degrees (Federmeier and Kutas, 1999a, 1999b) as well as category-general expectations distributed over many word forms that share a family resemblance (Dikker et al., 2009).

The strength of expectations formed during language processing must be balanced against the potential neural cost of making a false prediction. Accordingly, expectations are graded or probabilistic in nature: different words may violate them to different degrees. We suggest that N400/late positivity patterns can be explained by the relationship between the strength/specificity of linguistic expectancy and the degree to which it is violated by a critical word. If a critical word partially satisfies a relatively weak/general expectation (such as the expectation for a breakable object in *He was worried he had broken his collection*), then the result is an N400, reflecting the effort required to access that word’s semantic information, but no late positivity, as the situation model is
not constrained enough to form specific expectations. This N400 can be reduced by the usual word-level and contextual factors, such as high frequency, or semantic association with recently encountered words or active event knowledge (as in the control sentence, He was worried he had broken his arm…).

However, if the critical word defies a relatively strong/specific expectation (such as the expectation for blanket in the high-constraint, low-cloze sentence, He was cold most of the night and finally got up to get another log.), then the result should be a late positivity – in this case a frontal positivity – with the amplitude of the N400 dependent on the same factors mentioned above (examples from Federmeier et al., 2007). In this case, the comprehender has detected a strong conflict between the current word and the evolving model of the sentence’s meaning, and must revise the extant interpretation. An N400 and a late positivity can occur together when the critical word defies an extremely strong or specific expectation (for instance, when contextual constraint is near 1.0) and when the lexical and local associative factors that reduce N400s are not present (at least not to a greater degree than for the expected control word). In these cases, expectancy has become pre-activation. The word elicits an N400 because its meaning is effortful to activate due to poor contextual support (Kutas and Federmeier, 2000). The late positivity that follows (Federmeier et al., 2007) indexes the costs of suppressing the preactivated material and revising the failed sentence interpretation.

The view outlined above is highly compatible with the monitoring hypothesis, a contemporary account of the P600. The monitoring hypothesis views linguistic P600s as indexing a process of executive control, whereby the parsing and interpretation of input is continually being checked for errors based on contextually driven expectations (Van de Meeren-donk et al., 2009). In keeping with this hypothesis, Van de Meeren-donk et al. (2010) reported N400 responses to implausible words exhibiting mild conflict with contextual expectations (e.g., “The eye consisting of among other things a pupil, iris, and eyebrow….”), but enhancements of both N400 and a parietal late positivity for implausible words exhibiting stronger conflict (e.g., “The eye consisting of among other things a pupil, iris, and sticker….”). Such a strong conflict between what is expected and what is perceived can trigger a process of reanalysis to check for perceptual errors, a process that is indexed by P600s and perhaps by other late positivities.

Our account of the frontal positivity, and of the late positivities elicited by novel figurative language, is also compatible with the context updating hypothesis, which links linguistic P600s to the domain-general P300 response (Donchin, 1981; Pribram and McGuinness, 1975). The context updating hypothesis contends the P300 complex is best seen as a neural response to strategically important information that requires, “the model [to be] revised by building novel representations through the incorporation of incoming data into schema [a complex model of the ongoing environment] based on long-term memory data” (Donchin, 1981, p.508). In the case of the late positivity to language stimuli, Donchin’s schema should be construed as a detailed, slowly decaying, probabilistic model of the meaning and syntactic structure of the current linguistic context (Coulson et al., 1998). As outlined above, elicitation of the frontal positivity is likely related to domain-general processes of learning, memory, and executive control applied to the construction of meaning from linguistic input.

3.4. Conclusion

The present study compared the impact of predictability and semantic novelty/conventionality on ERPs elicited by sentence-final literal words. Unexpected but plausible completions of low constraint sentence fragments (LC and LN) elicited more negative N400 components and more positive frontal positivities than did expected completions of relatively high constraint sentence fragments (HC). Our manipulation of the novelty/conventionality of the semantic mapping (LC vs. LN) modulated the amplitude of the late portion of this frontal positivity elicited in the predictability manipulation, as the LC condition elicited more positive ERPs 700–900 ms than did LN sentences. Given the relationship between the predictability and novelty/conventionality effects, we have presented an account of a class of late positivities based on linguistic expectancy. The data in the present study were explained by appealing to linguistic expectations born of the situation model constructed during real time comprehension. A larger frontal positivity was elicited by the novel literal condition because the target words in this condition were most likely to defy the expectations encoded in the comprehender’s situation model. The resulting need to revise the situation model to integrate the meaning of the newly encountered word evoked a frontal positivity.

4. Experimental procedures

4.1. Participants

18 UC San Diego undergraduates (10 women) participated in exchange for course credit. Six additional participants were excluded from analysis due to blocked channels and/or excessive movement artifacts. The mean age of the participants was 20 years, ranging from 18 to 22. All participants were right-handed, with a mean score of +0.935 on the Edinburgh Handedness Inventory (Oldfield, 1971). All participants had normal or corrected-to-normal vision, and none had a history of head injuries or psychiatric problems.

4.2. Materials

Materials consisted of 240 triplets of sentences, with each member of a triplet ending in the same target word and belonging to a different experimental condition. Thus, each target word appeared 3 times: once in a high-cloze conventional mapping sentence (HC), once in a low-cloze conventional (LC) mapping context, and once in a low-cloze novel mapping context (LN). In addition, a true/false question was created for each stimulus sentence, to test comprehension and encourage semantic processing of the stimuli. For sample stimuli and comprehension questions, see Table 1.

Several norming surveys were conducted through an online applet in order to balance the stimuli between lists and conditions for a variety of statistical factors. None of the norming studies shared participants, nor did any norming
participants take part in the ERP experiment. The LC and LN conditions were a subset of materials employed by Coulson and Van Petten (2002), and thus were matched for cloze probability of the sentence final (target) word. HC sentence contexts were designed to promote reader’s expectation for target words employed in the two low-cloze conditions. The HC sentences were normed for cloze probability in an online norming study by presenting them – missing the final word – to 128 UCSD undergraduates who had to supply a one-word completion. The cloze probability means and standard deviations were 0.67 (SD=0.26) for the HC condition, 0.09 (SD=0.16) for the LC condition, and 0.08 (SD=0.13) for the LN condition. LC and LN sentences were also matched for the semantic distance between sentence-final words and their preceding sentence context, as assessed by latent semantic analysis (LSA).

All stimuli including the target words were rated for sentence-level concreteness in a separate study: three groups of 41, 37 and 37 UCSD students each were given 240 complete sentence stimuli (80 from each condition) and asked to rate them for concreteness on a scale of 1–5, with 1 being the most concrete and 5 the most abstract. Concreteness ratings for target words alone were obtained in the same way in yet another norming study of 30 participants. Sentence-level concreteness ratings did differ between conditions by a small but significant margin. Because each target word occurred in each condition, the conditions were matched for all word-level factors.

The 240 sentence triplets were split among 3 lists, ensuring that each final word appeared exactly once per list, and each list contained 80 sentences in each of the 3 conditions. With this constraint in mind, the sentences were distributed to ensure that within each list, the conditions matched as closely as possible on the following factors: mean length of final word (M=6.4, SD=2.0), mean number of orthographic neighbors of final word (M=4.2, SD=2.8), mean overall frequency of final word, per million words (M=86.1, SD=8.1), and word concreteness (M=2.1, SD=1.3). The LC and LN conditions were also matched within lists for the sentence-level factors of cloze probability and relatedness of context to target word, as assessed by LSA. Table 4 summarizes the statistical properties of the stimuli.

### 4.2.1. Semantic mapping types

In order to investigate differences in brain activity contingent on the type of semantic mapping being calculated, the 240 LN stimuli were divided into categories post-hoc based on the type of semantic mapping in the situation described. Because a large number of stimuli (around forty per condition) are generally needed to compare ERPs, the categories were necessarily somewhat general. As a first step, the LN stimuli were exhaustively divided into categories based on the relationship between the target word and the previously mentioned concept onto which its meaning is mapped. In the list below, the number of stimuli in each category is given in parentheses.

- Deception (21): To impress the guests, the duke pretended to be a prince.
- Mistake (64): The patient in room 203 has gotten the impression the doctors are masseurs.
- Symbolism (10): In the cartoon, it’s pretty clear that the outstretched tongue means thirst.

### Table 4 – Values for stimulus properties and their standard deviations in parentheses.

<table>
<thead>
<tr>
<th>Property</th>
<th>High-cloze conventional</th>
<th>Low-cloze conventional</th>
<th>Low-cloze novel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word length</td>
<td>6.4 (2.0)</td>
<td>6.4 (2.0)</td>
<td>6.4 (2.0)</td>
</tr>
<tr>
<td>Concreteness</td>
<td>2.1 (1.3)</td>
<td>2.1 (1.3)</td>
<td>2.1 (1.3)</td>
</tr>
<tr>
<td>Orthographic neighbors</td>
<td>4.2 (2.8)</td>
<td>4.2 (2.8)</td>
<td>4.2 (2.8)</td>
</tr>
<tr>
<td>Frequency per million words</td>
<td>86.1 (8.1)</td>
<td>86.1 (8.1)</td>
<td>86.1 (8.1)</td>
</tr>
<tr>
<td>Sentence-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latent semantic analysis</td>
<td>0.33 (0.15)</td>
<td>0.23 (0.12)</td>
<td>0.23 (0.11)</td>
</tr>
<tr>
<td>Concreteness</td>
<td>2.11 (0.31)</td>
<td>2.28 (0.3)</td>
<td>2.43 (0.33)</td>
</tr>
<tr>
<td>Cloze probability</td>
<td>0.67 (0.26)</td>
<td>0.09 (0.16)</td>
<td>0.08 (0.13)</td>
</tr>
</tbody>
</table>

- Pretend (41): The kids used the treehouse as a palace.
- Transformation (16): At one time, this movie theater was a cathedral.
- Substitution (60): To fix the clock, he substituted a paper clip for a spring.
- Repurposing (28): Yesterday, she used a spoon as a mirror.

From these categories, two larger categories were created. The “Mental” category comprised Deception, Mistake, Symbolism and Pretend, while the “Physical” category comprised Transformation, Substitution, and Repurposing. However, these larger categories were not matched in terms of any of the statistical factors controlled for in the group of stimuli as a whole. For the purposes of statistical comparisons between the two categories, we therefore removed a small subset of the stimuli in each, in order to balance them for cloze probability (Mmental=0.079, Mphysical=0.072, t=0.4, p=0.68), target word frequency (Mmental=85.8, Mphysical=87.1, t=1.2, p=0.24), target word length (Mmental=6.4, Mphysical=6.28, t=0.45, p=0.66), and target word’s number of orthographic neighbors (Mmental=2.79, Mphysical=2.79, t<0.01, p=0.99) and semantic distance from the sentence context as measured by LSA cosine (Mmental=0.23, Mphysical=0.22, t=0.8, p=0.41). After equalizing the Mental and Physical categories for these factors, they contained 115 and 99 stimuli, respectively. However, sentence-level concreteness still differed significantly, with the sentences in the Mental category rated as slightly more abstract by norming participants (Mmental=2.51, Mphysical=2.36, t=3.3, p=0.001).

### 4.3. Procedure

All stimuli were presented in white letters in Arial font on a black background, while participants sat 40 in. away from the monitor. A small orange fixation dot was present at the center of the monitor at all times. The beginning of each trial was signaled by a fixation cross in the center of the monitor.

1 All t-tests were 2-tailed and unpaired.

Please cite this article as: Davenport, T., Coulson, S., Predictability and novelty in literal language comprehension: An ERP study, Brain Res. (2011), doi:10.1016/j.brainres.2011.07.039
time. Each word was presented centrally for 200 ms, with a 300-ms inter-stimulus interval.

1000 ms after the disappearance of the final word, the word READ? was presented in green letters. The participant was instructed to indicate with a button press whether they had successfully read the entire stimulus sentence. After another 1000-ms delay, the comprehension question was presented for 6 s, during which time the participant answered true or false with a button press. 1000 ms after the comprehension question disappeared, a fixation cross appeared, signaling the start of the next trial. To minimize movement artifacts, participants were instructed to refrain from blinks or eye movements at all times except during the presentation of the comprehension question. Participants were given feedback about their blinking behavior, and they were allowed a short break after each block of 20 trials. Total time in the testing chamber was about sixty-five minutes.

4.4. EEG recording and analysis

EEG was recorded with 29 tin electrodes in an Electro-Cap mesh cap, organized in the International 10–20 configuration. Recordings were taken from 8 lateral sites: T5/6, T7/8, FT 7/8, F7/8, 10 medial sites: F3/4, C3/4, Cz, FC 3/4, and F3/4; 5 midline sites: Pz, Cpz, Cz, FCz, and Fz; 3 frontal sites: FP1/2 and FPz; and 3 occipital sites: O1/2 and Oz. Three additional electrodes were placed at the outer canthi of the eyes and below the left eye, to record eye movements and blinks. At all sites, electrical impedance was reduced below 5 kΩ with gentle abrasion. All EEG recorded was referenced on-line to a single electrode on the left mastoid, and later re-referenced to an average of the left and right mastoid electrodes. The EEG was amplified using a SA Instrumentation bioelectric amplifier, digitized online at 250 Hz.

ERPs were timelocked to the last word of each stimulus sentence. ERPs were examined for artifacts due to movement and channel blockage: overall, about 17% of trials were rejected for containing artifacts (MHC = 17.9%, MLC = 18.0%, MN = 16.5%); there was no effect of sentence type, indicating that trials were equally likely to be rejected, regardless of type (F[2,34] = 0.7, p = 0.48). Subsequent to artifact rejection, ERPs were averaged within conditions.

Only stimuli which participants indicated they were able to read were included in the analysis. ERP amplitudes were compared to a 100-ms pre-stimulus baseline. Four temporal windows were of primary interest: the P2 window (200–300 ms), the N400 window (300–500 ms), and two post-N400 windows to be analyzed for late positivity amplitudes (500–700 ms and 700–900 ms). Averaged ERP amplitudes were analyzed with several repeated measures ANOVA with factors sentence type (HC, LC, and LN) and electrode site (29 levels) as the independent variable. Planned comparisons were also conducted using factors Conventionality (LC/LN) and electrode site (29 levels). Post hoc analyses included topographic factors such as Hemisphere (left/right) and Anterior/Posterior.

Unless otherwise specified, all ANOVA’s were performed on mean amplitude measures. To compensate for the effect of repeated measures, the Greenhouse–Geisser correction was applied to the results of all ANOVA’s done on ERP data (Greenhouse and Geisser, 1959). All p-values cited are the corrected ones.

REFERENCES


