

Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension

Megan A. Boudewyn · Debra L. Long · Tamara Y. Swaab

© Psychonomic Society, Inc. 2015

Abstract The goal of this study was to investigate the use of the local and global contexts for incoming words during listening comprehension. Local context was manipulated by presenting a target noun (e.g., “cake,” “veggies”) that was preceded by a word that described a prototypical or atypical feature of the noun (e.g., “sweet,” “healthy”). Global context was manipulated by presenting the noun in a scenario that was consistent or inconsistent with the critical noun (e.g., a birthday party). Event-related potentials (ERPs) were examined at the feature word and at the critical noun. An N400 effect was found at the feature word, reflecting the effect of compatibility with the global context. Global predictability and the local feature word consistency interacted at the critical noun: A larger N200 was found to nouns that mismatched predictions when the context was maximally constraining, relative to nouns in the other conditions. A graded N400 response was observed at the critical noun, modulated by global predictability and feature consistency. Finally, post-N400 positivity effects of context updating were observed to nouns that were supported by one contextual cue (global/local) but were unsupported by the other. These results indicate that (1) incoming words that are compatible with context-based expectations receive a processing benefit; (2) when the context is sufficiently constraining, specific lexical items may be activated; and (3) listeners dynamically adjust their expectations when input is inconsistent with their predictions, provided that the inconsistency has some level of support from either the global or the local context.

Keywords Prediction · Discourse · Semantics · ERPs

M. A. Boudewyn (✉) · D. L. Long · T. Y. Swaab
University of California, Davis, California
e-mail: maboudewyn@ucdavis.edu

Intruction

It has long been recognized that contextual constraint has a rapid facilitatory effect on the processing of incoming words. The processing of words in reading and listening is modulated by the situations that are described in discourse (global context) and the meanings of adjacent words (local context) (Boudewyn, Gordon, Long, Polse, & Swaab, 2012; Camblin, Gordon, & Swaab, 2007; Federmeier & Kutas, 1999; van Berkum, Hagoort, & Brown, 1999; van Berkum, Zwitserlood, Hagoort, & Brown, 2003; Van Petten, 1993; Van Petten, Weckerly, McIsaac, & Kutas, 1997). However, the role of prediction in language processing has been a matter of considerable debate. On the one hand, the nearly unlimited number of possible continuations to any sentence has led some theorists to argue that predictive mechanisms would be impractical and unlikely to be successful (Forster, 1981; Jackendoff, 2002). On the other hand, intuition and anecdotal experience would suggest that language is often predictable, enabling us to complete a conversational partner’s sentences or to anticipate an event in a story. In addition, numerous studies have suggested that individuals can make predictions about language input during comprehension under certain circumstances (Brothers, Swaab, & Traxler, 2015; DeLong, Urbach, & Kutas, 2005; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004; see Lau, Holcomb, & Kuperberg, 2013, for effects of prediction in a word-priming paradigm). It seems that these predictions can be made at various levels, including the activation of upcoming events and event structure, semantic features, or specific lexical forms (Kuperberg, 2013). In the present study, we focus on the possibility that prediction plays an important role in processing of the semantic features and lexical forms of incoming words.

This study had three main goals. First, we used event-related potentials (ERPs) and a story-listening paradigm in

order to examine semantic feature prediction during comprehension. Second, we investigated the contributions of two “levels” of context (the global, discourse context and the local, “prime word” context) to two “levels” of prediction: the activation of semantic features and the prediction of specific word forms. Finally, we examined the costs that may be associated with disconfirmed predictions. We introduce each of these issues in turn.

Semantic feature prediction

Evidence for activation of the semantic features of incoming words has come from several ERP studies that have shown that the processing of incoming words in context is facilitated when words that are unpredictable (or even incongruent) in the context share some semantic features with predictable words (Federmeier, 2007; Federmeier & Kutas, 1999; Federmeier, Kutas, & Schul, 2010; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Otten & Van Berkum, 2007; Paczynski & Kuperberg, 2012). For example, readers have shown a graded N400¹ effect as a function of semantic feature overlap with predictable words (e.g., Federmeier & Kutas, 1999). As compared to predictable words (e.g., “palms” in the context of “They wanted the hotel to look more like a tropical resort. Along the driveway they planted...”), unexpected but within-category words (e.g., “pines”) elicit a smaller N400 than do unexpected different-category words (e.g., “tulips”) (e.g., Federmeier & Kutas, 1999). In another study, readers showed a reduced N400 for event-related words (e.g., “luggage” in the context of a “travel” scenario) as compared to event-unrelated words (e.g., “vegetables”) even when both were implausible (and importantly, unpredictable), given their placement within a particular sentence (e.g., “It can take several hours to get through luggage/vegetables”) (Metusalem et al., 2012; see Nieuwland & Van Berkum, 2005, and Paczynski & Kuperberg, 2012, for related manipulations resulting in modulations of the N400 for scenario-related but lexically unpredictable words). The anticipation of semantic features was probed in the present study by examining the ERP response to words that were not themselves predictable but that represented semantic features of highly predictable nouns that were presented farther downstream. We constructed two-sentence stories in which a word in the second sentence (target noun) was highly predictable (e.g., “cake”) and was preceded by a feature word that was a typical or atypical feature of the critical noun (“sweet” or “healthy”), as in Example(a) below. See Table 1 for additional examples.

¹ The N400 is a negative-going ERP that is reduced in amplitude as a function of semantic fit and lexical properties (see Kutas & Federmeier, 2000, 2011, for reviews; Swaab, Ledoux, Camblin, & Boudewyn, 2012).

- (a) Frank was throwing a birthday party, and he had made the dessert from scratch. After everyone sang, he sliced up some...
... sweet/healthy and tasty cake/veggies that looked delicious.

If semantic features are activated given a constraining discourse context, then the feature words (sweet/healthy) should show an N400 effect of feature compatibility, even though the feature words are unpredictable. Specifically, if predictions about the upcoming input include context-compatible semantic features, then compatible-feature words such as “sweet” should result in a reduced N400 relative to incompatible-feature words such as “healthy.”

Predictability based on global and local context

A second goal of the study was to examine the extents to which the global, discourse context and the local, word-level context contribute to two types of predictions: lexical prediction and semantic feature prediction. *Lexical prediction* refers to the anticipation of a specific word, including its form, prior to encountering that word. *Semantic feature prediction* refers to a broader category of expectations about the semantic information that is likely to be encountered, but falls short of a

Table 1 Sample stimuli across the four conditions

Context: Frank was throwing a birthday party, and he had made the dessert from scratch. After everyone sang, he sliced up some...
Globally Predictable/Locally Consistent: ... <u>sweet</u> and tasty <u>cake</u> that looked delicious.
Globally Predictable/Locally Inconsistent: ... <u>healthy</u> and tasty <u>cake</u> that looked delicious.
Globally Unpredictable/Locally Consistent: ... <u>healthy</u> and tasty <u>veggies</u> that looked delicious.
Globally Unpredictable/Locally Inconsistent: ... <u>sweet</u> and tasty <u>veggies</u> that looked delicious.
True/False: Frank planned to make hamburgers for the party.
Context: In the summertime, Ron loves to barbecue and drink Coronas, so he always keeps some beer and citrus to go with it in the fridge. His drink is not complete unless he has some...
Globally Predictable/Locally Consistent: ... <u>slices</u> of refreshing <u>lime</u> in it.
Globally Predictable/Locally Inconsistent: ... <u>cubes</u> of refreshing <u>lime</u> in it.
Globally Unpredictable/Locally Consistent: ... <u>cubes</u> of refreshing <u>ice</u> in it.
Globally Unpredictable/Locally Inconsistent: ... <u>slices</u> of refreshing <u>ice</u> in it.
True/False: Ron likes to barbecue.

Critical words are underlined.

specific, word-form prediction. As was discussed in the previous section, the results of several studies have indicated that readers/listeners activate semantic features during comprehension (Federmeier, 2007; Federmeier & Kutas, 1999; Federmeier et al., 2010; Federmeier et al., 2007; Metusalem et al., 2012; Nieuwland & Van Berkum, 2005; Otten & Van Berkum, 2007; Paczynski & Kuperberg, 2012). Recent evidence has also suggested that readers and listeners make specific lexical predictions under some circumstances (Brothers et al., 2015; DeLong et al., 2005; van Berkum et al., 2005; Wicha et al., 2004). For example, processing difficulties have been found when the phonological or syntactic markers of incoming words are inconsistent with those of predicted words (DeLong et al., 2005; Szewczyk & Schriefers, 2013; van Berkum et al., 2005; Wicha et al., 2004).

In the present study, we addressed the following question: When the discourse context is highly constraining, how do global and local sources of contextual information interact to influence expectations about the upcoming input? Global, discourse context appears to exert a powerful influence on the processing of incoming words (even when the context is not constraining), whereas the local context immediately preceding an incoming word appears to exert a relatively weaker influence on processing (Boudewyn et al., 2012; Camblin et al., 2007; although there is substantial variability in the relative weighting of contextual information—e.g., Boudewyn, Long, & Swaab, 2013). In constraining contexts, lexical predictions based on the global discourse message may “trump” local inconsistencies: For example, a strong expectation for “cake” that has developed over the course of a birthday scenario may not be affected by the presence of a lone incompatible semantic feature, such as “healthy.”

To investigate this issue, we examined nouns (“cake/veggies,” in the example above) that followed semantic feature words (“sweet/healthy”) in the second sentence of each story context. If lexical predictions that are based on global context are robust enough to withstand brief, local inconsistencies, then an N400 response to nouns in the “healthy and tasty cake” condition should be reduced to the same extent as in the “sweet and tasty cake” condition. However, the presence of conflicting local context may serve to weaken predictions that are based on the global, discourse context. In that case, the processing of nouns in the “healthy and tasty cake” condition should not receive the same level of facilitation as the processing of nouns in the “sweet and tasty cake” condition, resulting in an N400 effect when comparing the two. Likewise, a discourse-incompatible, but locally consistent, feature should facilitate the processing of nouns in the “healthy and tasty veggies” condition, relative to the “sweet and tasty veggies” condition. This pattern of results would indicate that listeners dynamically update their expectations, even when the global context is highly constraining, and importantly, that they do so

quickly after receiving just one word of conflicting local context.

Disconfirmed predictions

The third goal of this study was to investigate the processing costs that may be associated with disconfirmed predictions, and the extent to which they are triggered by prediction-inconsistent global and local cues. If readers and listeners activate upcoming semantic features and/or specific lexical items, then encountering input that is either partially or fully at odds with those expectations should have an impact on processing. We suggest that the impact of unexpected input on processing may take several forms, depending on the strength of the prediction and the potential for integration and updating of expectations based on the unexpected input.

First, a strong prediction for a specific word form that is disconfirmed should have an early influence on the processing associated with detection of the mismatch between the expected and received forms. Indeed, ERP differences have been observed within a few hundred milliseconds of encountering an incoming word in situations in which it is possible to “diagnose” a mismatch between the expected and received input relatively early, as is the case in syntactic category violations (Lau, Stroud, Plesch, & Phillips, 2006; van Berkum et al., 2005). Early effects (N200s) have also been observed in spoken language when phonemic input has not matched the most predictable word form² (Connolly & Phillips, 1994; Diaz & Swaab, 2007; van den Brink, Brown, & Hagoort, 2001; see also Brothers et al., 2015, for a similar effect during reading).

In the present study, an early effect (N200) should be observed at the critical nouns in the second sentence of the story contexts (*cake/veggies*, in the example above) if listeners make specific lexical predictions (that include a word form). Namely, nouns that do not match the most expected word form should elicit a larger N200 than would nouns that are consistent with the most expected word form (i.e., the word form *veggies* is an easily detected mismatch to the most globally predicted word, *cake*). However, if listeners are quickly influenced by the immediately preceding (local) context as well as by the global, discourse context, then we predict that the global and local context effects will interact. Only the most constraining condition (when both the global and local contexts are aligned to predict the same item) will elicit an early

² These early effects might be characterized as indicative of the cost associated with processing the unexpected versus the expected form, but it would also be reasonable to characterize these effects as indicative of the “lack of benefit” that the unexpected forms received relative to the expected forms.

effect, such that a larger early negativity will be observed for nouns that mismatch the prediction (sweet and tasty VEGGIES) than for nouns that match the prediction (sweet and tasty CAKE).

Second, disconfirmed predictions may also lead to later processing costs that are associated with updating the developing discourse representation to accommodate the unexpected input. This type of cost has been an issue of considerable interest following the recent observation of late positive ERP deflections in response to unexpected words in context (Van Petten & Luka, 2012). These late effects, termed *post-N400 positivities* (PNPs), have been interpreted as reflecting the costs of encountering input that is unexpected, but still plausible, although the precise nature of the costs has yet to be determined. One possibility is that the PNP reflects inhibitory processes that are associated with suppressing the predicted, but not encountered, word (Kutas, 1993). Another (non-mutually-exclusive) explanation is that the PNP reflects processing that is required in order to discard previous expectations and update representations after receiving input that is plausible, but not consistent with, a prediction (Kuperberg, 2013). If global and local contexts combine to influence the strength of predictions about the upcoming input, then, according to either account, a PNP should be elicited by the globally unpredictable nouns following locally inconsistent feature words (sweet and tasty VEGGIES), since this is the condition in which listeners are most likely to predict a specific lexical item. A PNP should also be elicited by the other two conditions that contain a conflicting cue (healthy and tasty CAKE, and healthy and tasty VEGGIES), because they entail a violation of expectations that may require updating the discourse representation, as well.

Thus, several ERP effects have been identified as possible signatures of predictive processing, most notably: (1) early effects (such as the N200), which may reflect the initial detection of a disconfirmed prediction; (2) the N400, which is reduced as a function of contextual fit; and (3) late effects in the PNP window, which is the least well-characterized effect to date, but which appears to be related to costs that are associated with disconfirmed predictions. In the present study, we used these ERP effects to investigate our three main questions of interest about prediction during listening comprehension.

Method

Participants

Twenty undergraduates (15 female, five male) from the University of California, Davis, gave informed consent and participated for course credit. All were right-handed, native English speakers, with no reported problems with hearing/

reading nor a history of neurological/psychological disorders (average age = 18.63, range = 18–20).

Materials

The materials consisted of 132 two-sentence stories, in which two words in the second sentence were manipulated: a feature word, and a critical noun that occurred a few words later. The ERPs were time-locked to the feature words (two conditions) and to the critical nouns (four conditions). To create the stories, two variables were orthogonally manipulated: global predictability and local consistency. This resulted in four conditions: globally predictable/locally consistent, globally predictable/locally inconsistent, globally unpredictable/locally consistent, and globally unpredictable/locally inconsistent (see Table 1 for examples).

The critical words never appeared in a sentence-final position. Stories were divided into four lists and counterbalanced such that each list contained an equal number of stories in each condition and only one condition per set was included in each list. In addition, 120 filler stories were included; a subset of these (40 stories) included either congruent or anomalous words midway through the second sentence. These were included in order to minimize the salience of the experimental manipulation. Globally predictable nouns were highly predictable midway through the second sentence of each story, and globally unpredictable nouns were unexpected at that same point. This was established by a norming study, in which 45 participants (who did not participate in the ERP study) received the stories with the second sentence truncated before the highly predictable noun. They were asked to generate a word that best completed the story. The results showed that each story could be completed at that point by a highly predictable noun (cloze probability = 78 %). The unpredictable nouns were semantically and syntactically possible continuations of the story but had a cloze probability of 0 %. Critically, neither the globally predictable nor the globally unpredictable nouns were ever presented at that point in the story; instead, a feature word (e.g., sweet) replaced the noun. The feature word was either locally consistent with the globally predictable noun or locally consistent with the globally unpredictable noun. All of the feature words were unexpected at that point, whether the norming participants were asked to provide a single-word continuation (cloze probability = 0.01 %) or a multiple-word continuation (cloze probability = 1.81 %).

Because the unpredictability of the feature words was critical to our experimental manipulation, we also conducted a modified cloze test in order to rule out the possibility that repeated exposure to passages containing feature words that were followed by nouns would lead participants to predict the feature words. Sixty participants who did not participate in the previous norming tests or in the ERP study were asked to

complete the stories, with no constraints on the number of words allowed for the continuations. Four lists were created, such that in a given list, 25 % of the experimental stories were truncated at the point at which the critical nouns were highly predictable and the feature words were unpredictable; the rest of the experimental stories were truncated prior to the final word in the second sentence (i.e., after the feature word/noun combination). We included filler stories so that a high proportion of the stories in a list would contain feature word/noun combinations (85 %), in order to maximize the opportunity for participants to pick up on this combination and begin to predict the feature words themselves. Despite this list construction, the results confirmed that following repeated exposure to stories containing feature word/noun combinations, all feature words were lexically unexpected (cloze probability = 0.01 %).

A feature word described a prototypical semantic feature of the predictable or unpredictable noun, and the critical noun was moved to a position at least one word after the unpredictable feature word (average = 2.2, range = 1–6). Feature words were either selected from published lists of feature norms (McRae, Cree, Seidenberg, & McNorgan, 2005) or generated for this study, and they were verified as characteristic of the nouns in a separate norming study ($n = 60$). Participants were asked to rate how characteristic each feature word was of its noun on a 7-point scale (1 = *very characteristic*, 7 = *very uncharacteristic*). Compatible feature words were rated as highly characteristic of the predictable nouns (average = 1.94, range = 1–3.73) and highly uncharacteristic of the unpredictable nouns (average = 6.14, range = 4.06–7) ($p < .001$). In contrast, incompatible feature words were rated as highly characteristic of the unpredictable nouns (average = 1.86, range = 1–3.86) and highly uncharacteristic of the predictable nouns (average = 5.86, range = 4–7; $p < .001$).

Feature words were matched on length (compatible: average = 6.37, range = 3–13; incompatible: average = 6.28, range = 3–12; $p = .72$) and frequency (compatible: average = 2.25, range = 0–3.8; incompatible: average = 2.17, range = 0–3.9; $p = .46$), using the LG10CD measure from the SUBTLEXus database (<http://expsy.ugent.be/subtlexus/>). The feature words were also matched for number of syllables (compatible: average = 1.75, range = 1–5; incompatible: average = 1.78, range = 1–4) and number of phonemes (compatible: average = 4.49, range = 2–9; incompatible: average = 4.5, range = 2–10), using the MRC Psycholinguistic Database (<http://websites.psychology.uwa.edu.au/school/MRCDatabase/>). All stimuli were recorded for spoken presentation and were then matched on spoken duration (milliseconds). This resulted in the exclusion of 15 items. The final stimulus set on which all analyses were performed contained 117 items (average duration: compatible, 484.04, range = 282–983; incompatible, 505.6, range = 240–879; $p = .2$). Nouns were matched on the same parameters: length (predictable: average = 5.86, range = 3–13;

unpredictable: average = 6.16, range = 3–12; $p = .2$), frequency (predictable: average = 2.26, range = 0–3.78; unpredictable: average = 2.19, range = 0–3.9; $p = .42$), number of syllables (predictable: average = 1.77, range = 1–4; unpredictable: average = 1.71; range = 1–4), number of phonemes (predictable: average = 4.47, range = 2–8; unpredictable: average = 4.44, range = 2–8) and spoken duration (predictable: average = 509.65, range = 301–780; unpredictable: average = 510.78, range = 253–813; $p = .94$).

All stimuli were spoken by a female with a natural inflection and speaking rate. Stimuli were recorded using a Schoeps MK2 microphone and Sound Devices USBPre A/D (44, 100 Hz, 16 bit) in a sound-attenuating recording booth. The acoustic onset and offset of each critical word was determined by visual inspection of the speech waveform and by listening, using speech-editing software (Audacity, by Soundforge). The context sentence (Sentence 1) and the critical sentence (Sentence 2) of each story were recorded separately. A 1-s silence was inserted between the two sentences using Presentation software, as has been done in previous work to approximate the duration of naturally produced pauses (Boudewyn et al., 2012). True/false comprehension questions were included after each story; these questions did not focus on the critical words. The correct answer to half of the questions was true.

Procedure

Participants listened to the stories while sitting in an electrically shielded, sound-attenuating booth. Stories were presented through Beyer dynamic headphones using the Presentation software (www.neurobs.com/). Trials began with a white fixation cross in 16-point Tahoma font in the center of a black screen situated 100 cm in front of participants. The fixation cross appeared on the screen 1,000 ms before story onset and remained visible on the screen throughout story presentation and for 1,000 ms after. The cross was then replaced by a comprehension question. Participants responded by pressing one of two keys on a keyboard, corresponding to “true” or “false” (left corner “Z” for true and right corner “M” for false). Comprehension questions remained on the screen until participants had made a response, which triggered the start of the next trial.

Participants were asked to remain alert and focused during the experiment and to fixate on the white fixation cross whenever it was present. They were instructed that they were free to make eye movements during the comprehension question portion of the trials. Each session was divided into 11 blocks to allow for short breaks. Event codes were sent out at the onset of the critical feature words and nouns and used for offline averaging of the electroencephalography (EEG) signal.

ERP recording and data reduction The EEG was recorded from 29 tin electrodes in an elastic cap (Electro-Cap International), and from electrodes on the outer canthi, below and above the left eye (to monitor eye movements and blinks), and on both mastoids. The right mastoid served as the recording reference, and the left was used for offline algebraic re-referencing (to the average of both mastoids) for all channels except the eye channels. The eye channels were referenced to each other (above-left to below-left, and left outer canthus to right outer canthus). The EEG signal was amplified with band-pass cutoffs at 0.01 and 30 Hz and was digitized online at a sampling rate of 250 Hz (Neuroscan Synamps 2). Impedances were kept below 5 Ω .

Data processing was performed using MATLAB, with the EEGLAB toolbox and ERPLab plugin (Lopez-Calderon & Luck, 2014) and custom MATLAB and UNIX routines. Independent component analysis (ICA) artifact correction was used to correct for eye blinks (participants were not instructed to refrain from blinking during the experiment). The single-trial waveforms were screened for amplifier blocking, muscle artifacts, and horizontal eye movements over epochs of 1,200 ms, starting 200 ms before the onset of the critical word. Average ERPs were computed over all artifact-free trials in the four conditions. All ERPs were filtered offline with a Gaussian low-pass filter with a 25-Hz half-amplitude cutoff. Excessive noise in the data from one participant resulted in that participant being excluded from all analyses, since too few artifact-free trials remained to achieve an acceptable signal-to-noise ratio after artifact removal.

Results

The participants were highly accurate on the comprehension questions, scoring an average of 94.38 % (range = 87.7 %–97.9 %).

ERPs to the feature words are displayed in Fig. 1; ERPs to the critical nouns are displayed in Fig. 2. Simple-effect comparisons for ERPs at the critical nouns are displayed in Fig. 3. Figure 4 depicts differences in the topographic distribution of the N200 and N400 responses to the critical nouns.

Separate repeated measures analyses of variance (rANOVAs) were conducted for the feature words and the critical nouns over the midline (Fz, Cz, Pz, POz), medial (FC1, FC2, C3, C4, CP1, CP2), and lateral (F3, F4, FC5, FC6, CP5, CP6, P3, P4) electrode columns. The within-subjects variable, feature compatibility (compatible, incompatible), was used in the rANOVA for the feature-continuation analysis. For the critical-noun analysis, the within-subjects variables, global predictability (predictable, unpredictable) and local consistency (consistent, inconsistent), were included. All rANOVAs included topographic distribution factors: For the midline column, this was Electrode

Feature Compatibility Effect

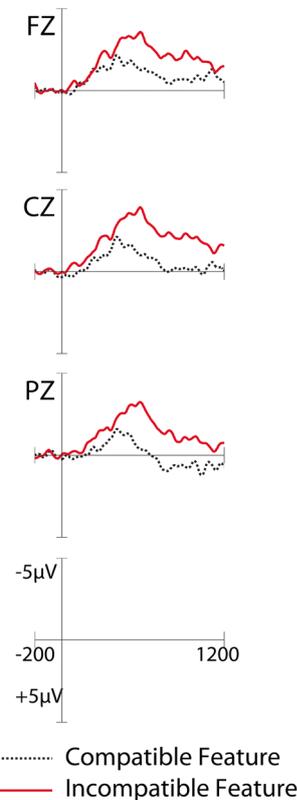


Fig. 1 Feature compatibility effect at the first time lock (feature words). Representative frontal, central, and posterior midline electrode sites are shown (Fz, Cz, and Pz, respectively). Negative is plotted up

Site (four levels), and for the medial and lateral analyses, these were Hemisphere (left vs. right) and Anteriority (medial [fronto-central, centro-parietal, parietal] vs. lateral [frontal, fronto-central, centro-parietal, parietal]). A Greenhouse–Geisser correction was used for *F* tests with more than one degree of freedom in the numerator. The rANOVAs were conducted on the mean amplitudes in three time windows: the N200 (200–300 ms), the N400 (300–600 ms), and the PNP (900–1,200 ms). Significant interactions in the critical-noun analysis were followed with simple-effect comparisons across the same electrode sites and time windows. The results are summarized below and are presented in full in Table 2 (feature words) and in Tables 3, 4, and 5 (critical nouns).

Feature words

We found no significant effect of feature compatibility at any electrode column in the N200 time window. In the N400 window, all electrode columns displayed a significant main effect of feature compatibility (midline, medial, lateral), such that compatible features showed a reduced negative deflection relative to incompatible features. In the PNP window, a significant main effect of feature compatibility emerged at the medial

Critical Nouns: Global Predictability & Local Consistency Effects

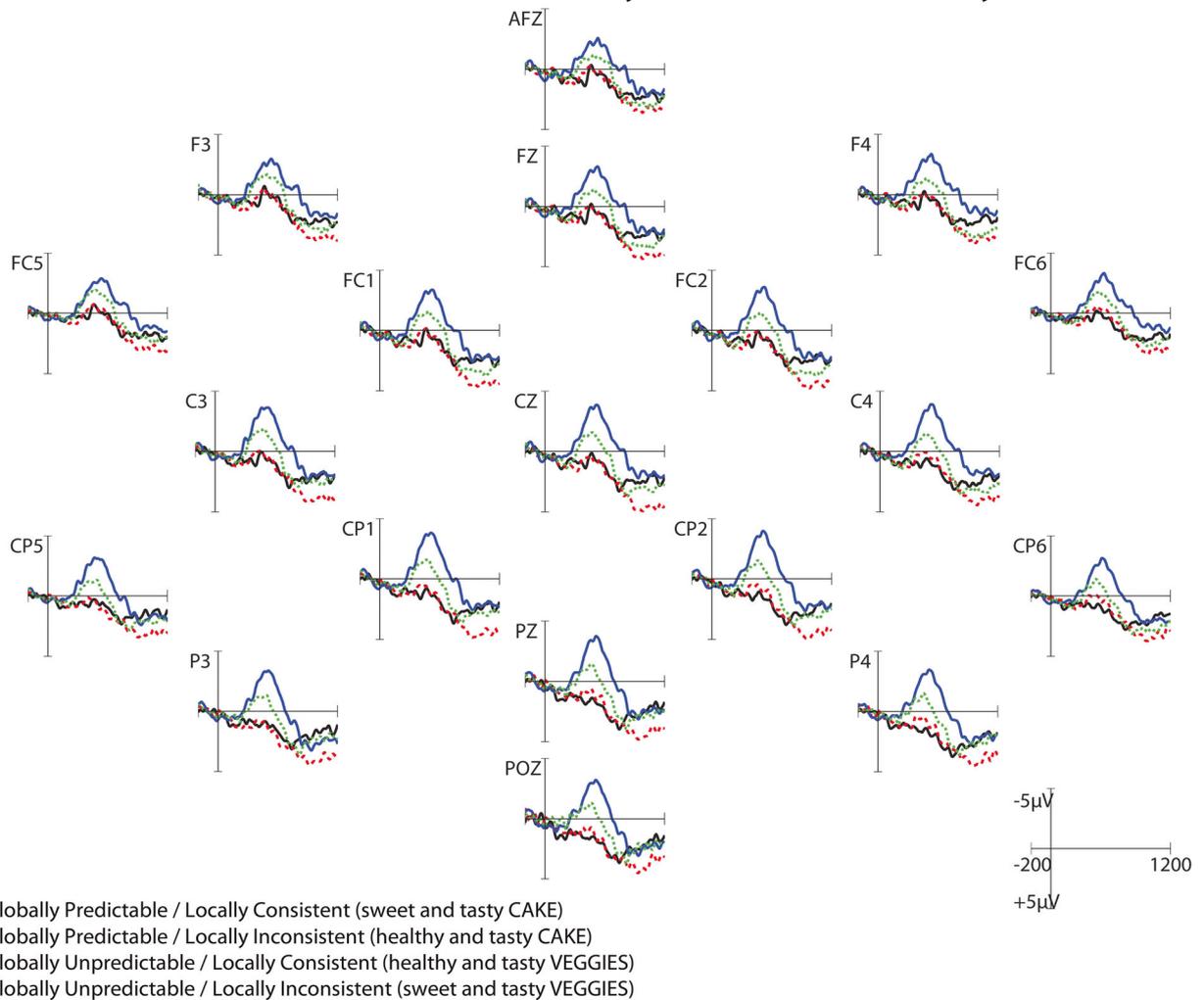


Fig. 2 Event-related potentials (ERPs) to the second time lock (critical nouns) in all four conditions. All electrodes tested are shown (midline, medial, and lateral columns). Negative is plotted up

and lateral columns, such that compatible features continued to show a reduced negative deflection relative to incompatible features.

Critical nouns

We observed significant effects of global predictability at the midline and lateral electrode columns in the N200 window, such that predictable nouns showed a reduced negative deflection relative to unpredictable nouns; for lateral sites, we also found a significant Global Predictability \times Anteriority \times Hemisphere interaction. There were no effects of local consistency. Global predictability and local consistency significantly interacted at both medial and lateral electrode sites. Follow-up analyses comparing locally consistent to locally inconsistent words showed (a) no effect for globally predictable words (sweet and tasty CAKE

vs. healthy and tasty CAKE) and (b) a Local Consistency \times Anteriority \times Hemisphere interaction for globally unpredictable words (healthy and tasty VEGGIES vs. sweet and tasty VEGGIES) at lateral electrode sites, such that locally consistent nouns showed a smaller negative deflection at right anterior electrode sites than did locally inconsistent nouns. Follow-up analyses comparing globally predictable to globally unpredictable words showed (a) a difference between locally consistent words (sweet and tasty CAKE vs. healthy and tasty VEGGIES) at the midline column, such that globally predictable words showed a smaller negative deflection than did globally unpredictable words, and (b) a difference between locally inconsistent words (healthy and tasty CAKE vs. sweet and tasty VEGGIES) at all columns, such that globally predictable words showed a smaller negative deflection than did globally unpredictable words.

Critical Nouns: Simple Effects Comparisons

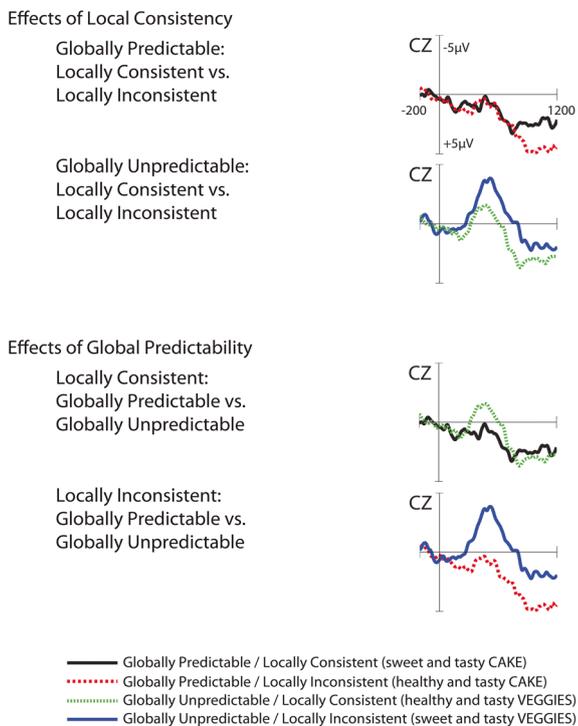


Fig. 3 Simple-effects comparisons at the second time lock (critical nouns). A representative central electrode site is shown. Negative is plotted up

In the N400, we observed significant effects of global predictability at all columns (midline, medial, lateral), such that predictable nouns showed a reduced negative deflection relative to unpredictable nouns. This effect was maximal at posterior electrodes, as can be seen in Fig. 2, and interacted with topographic factors (see Table 3). There were also significant effects of local consistency at all columns. Global predictability and local consistency interacted at medial and lateral electrode sites; at lateral sites, an additional Global Predictability \times Local Consistency \times Anteriority \times Hemisphere interaction was apparent. Follow-up analyses comparing locally consistent to locally inconsistent words showed (a) a difference between globally predictable words (sweet and tasty CAKE vs. healthy and tasty CAKE) at right posterior lateral electrode sites and (b) a difference between globally unpredictable words (healthy and tasty VEGGIES vs. sweet and tasty VEGGIES) at all electrode columns. Follow-up analyses comparing globally predictable to globally unpredictable words showed (a) a difference between the two locally consistent conditions (sweet and tasty CAKE vs. healthy and tasty VEGGIES) at all electrode columns and (b) a difference between the two locally inconsistent conditions (healthy and tasty CAKE vs. sweet and tasty VEGGIES) at all electrode columns.

In the PNP window, no significant effect of global predictability emerged at any electrode column, whereas the effect of local consistency significantly interacted with topographic factors at medial and lateral sites. Global predictability and local consistency significantly interacted at all electrode columns. Follow-up analyses comparing locally consistent to locally inconsistent words showed (a) a difference between the two globally predictable conditions (sweet and tasty CAKE vs. healthy and tasty CAKE) at all electrode columns, such that locally inconsistent words showed a larger positive deflection than did locally consistent words, and (b) a difference between the two globally unpredictable conditions (healthy and tasty VEGGIES vs. sweet and tasty VEGGIES) at lateral electrode sites, such that locally inconsistent nouns showed a greater positive deflection than did locally consistent nouns at right lateral sites. Follow-up analyses comparing globally predictable to globally unpredictable words showed (a) no difference between the two locally consistent conditions (sweet and tasty CAKE vs. healthy and tasty VEGGIES) at any electrode column and (b) a difference between the two locally inconsistent conditions (healthy and tasty CAKE vs. sweet and tasty VEGGIES) at all electrode columns, such that globally predictable words showed a greater positive deflection than did globally unpredictable words.

Topographic comparison of N200 and N400 As can be seen in Figs. 2 and 4 and in Tables 3, 4, and 5, there were substantial differences between the N200 and N400 effects at the critical noun, including differences in topographic distribution. In order to determine whether the differences in scalp distributions were statistically significant, a vector-scaling approach was adopted in which the size of the effects was normalized across participants (McCarthy & Wood, 1985; Ruchkin, Johnson, & Friedman, 1999; but see Urbach & Kutas, 2002). The analyses reported above showed that the N200 effect was driven primarily by a larger negative deflection to the globally unpredictable/locally inconsistent (sweet and tasty VEGGIES) condition than to all other conditions; this condition also showed the largest negative deflection in the N400 time window. The difference can be seen in Fig. 2. In order to determine whether responses in this condition were distinct in terms of distributions across the scalp, the vector-scaled difference between the globally unpredictable/locally inconsistent condition and the average of the three other conditions was used. This difference was the dependent measure in rANOVAs at the midline, medial, and lateral columns; time window was included as a within-subjects variable (N200, N400), along with the same topographic factors described above. Interactions of time window and the topographic factors would indicate significant differences in the distributions of the interaction across the N200 and N400 time windows.

Topographic Distribution Comparison: N200 vs. N400

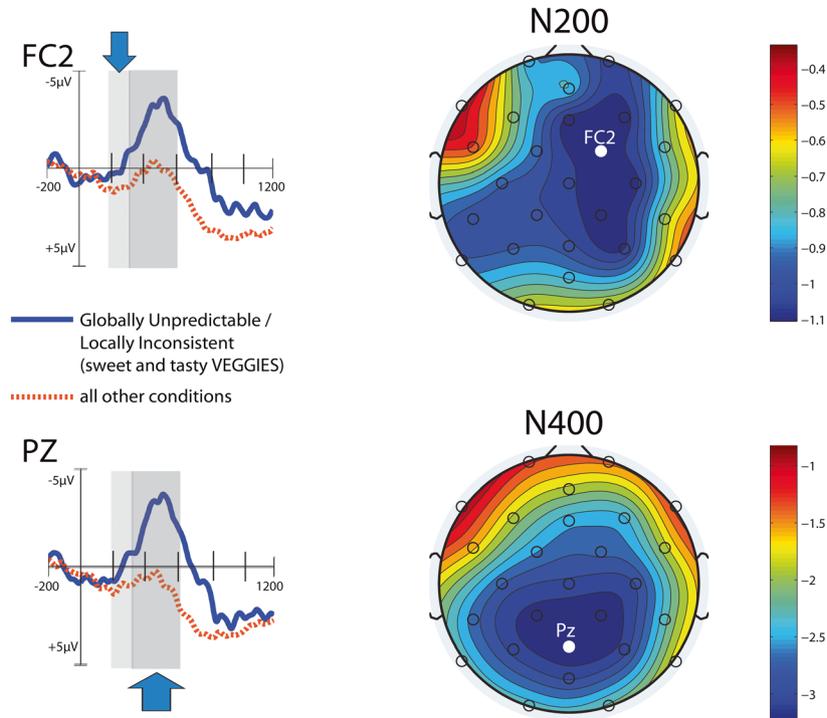


Fig. 4 Event-related potential (ERP) plots: The N200 and N400 effects for the globally unpredictable/locally inconsistent (sweet and tasty VEGGIES) condition, relative to the average of all other conditions, are shown at left. The frontal electrode at which the N200 effect was maximal (FC2) is shown at top left and highlighted with a white dot on the topographic map at top right. The electrode at which the N400 effect was maximal (Pz) is shown at bottom left, and highlighted with a white

dot on the topographic map at bottom right. The N200 window (200–300 ms) and N400 window (300–600 ms) are highlighted in shaded gray regions. Negative is plotted up. The topographic distributions of the globally unpredictable/locally inconsistent condition (sweet and tasty VEGGIES) minus all other conditions are shown at right; the distribution of the N200 effect is shown at top right, and the distribution of the N400 effect is shown at bottom right

Consistent with the topographic distribution of the effects depicted in Fig. 4, we found a significant Time Window \times Anteriority interaction at lateral electrode sites [$F(3, 54) = 3.702, p < .5$]. This reflects the more frontal distribution of the N200 effect, in contrast to the more posterior distribution of the N400 effect.

Discussion

Our goal in this study was to examine the influence of local and global contexts on prediction in listening comprehension. Specifically, we sought (1) to determine the extent to which the semantic features of predictable nouns were activated before the nouns appeared, (2) to investigate the interaction between global and local context on the prediction of semantic features and specific word forms, and (3) to examine processing costs that might be associated with encountering unexpected inputs. We examined ERPs at two points during the second sentence of short spoken stories. The first time lock

was to unpredictable words that were prototypical semantic features of upcoming, critical nouns. The second time lock was to the critical nouns themselves, which occurred a few words downstream from the feature words and were either highly predictable or unpredictable in the global context. Three ERP effects were of interest: the N200 (200–300 ms post-word-onset), N400 (300–600 ms post-word-onset), and PNP (following the N400; 900–1,200 ms post-word-onset). We discuss effects for each time lock and time window in turn.

Accessibility of semantic features prior to critical words

Previous studies have shown that processing is facilitated for words that share semantic features with those that are the best completion of a sentence (e.g., “pines” in a sentence in which “palms” is the best completion; Federmeier & Kutas, 1999). In the present study, we examined whether or not the processing of semantic features was facilitated for words that were unpredictable, given the preceding context, and that occurred *before* the critical (predictable) noun was heard. Since the feature words themselves were not predictable, we did not

Table 2 Event-related potential results for the feature words

	<i>df</i>	N200			N400			PNP		
		<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2
Midline										
Feature Compatibility	(1, 18)	3.56	–	–	12.54	**	.28	<1	–	–
Feature Compatibility × Electrode	(3, 54)	1.49	–	–	5.42	*	.01	<1	–	–
Medial										
Feature Compatibility	(1, 18)	2.22	–	–	13.2	**	.21	6.37	*	.20
Feature Compatibility × Hemisphere	(1, 18)	1.8	–	–	3.93	–	–	1.44	–	–
Feature Compatibility × Anteriority	(2, 36)	1.46	–	–	<1	–	–	<1	–	–
Feature Compatibility × Hemisphere × Anteriority	(2, 36)	2.78	–	–	<1	–	–	<1	–	–
Lateral										
Feature Compatibility	(1, 18)	1.71	–	–	10.99	**	.28	5.27	*	.12
Feature Compatibility × Hemisphere	(1, 18)	2.3	–	–	2.01	–	–	3.19	–	–
Feature Compatibility × Anteriority	(3, 54)	1.24	–	–	<1	–	–	<1	–	–
Feature Compatibility × Hemisphere × Anteriority	(3, 54)	<1	–	–	1.33	–	–	2.12	–	–

*** $p < .001$, ** $p < .01$, * $p < .05$, $^{\wedge}p < .07$

expect an N200 effect, which would have reflected a mismatch between the predicted word and the word that was actually presented. If, however, listeners activated the semantic features of the predictable nouns on the basis of the global context, then the N400 amplitude should be reduced for words that were prototypical features of globally predictable nouns, relative to those that were features of globally unpredictable nouns.

Consistent with our predictions, we observed no effect in the N200 window for the feature words. However, we did observe a significant N400 effect of feature compatibility, with compatible feature words showing a reduced N400 relative to incompatible feature words. This finding demonstrates that the semantic features of highly predictable words are accessible *before* the predictable word is heard. Our results are consistent with the current thinking about predictive processing during language comprehension, which posits that predictions about upcoming language input are made at different levels, including the semantic feature level (Kuperberg, 2013; Pickering & Garrod, 2013).

Prediction of lexical form

Our results showed that the amplitude of the N200 was significantly greater for critical words that were both globally unpredictable and locally unsupported than for all other conditions (e.g., sweet and tasty VEGGIES; see Table 1). Note that, in this condition, a specific lexical prediction (e.g., CAKE) was fully supported by all available context up until the critical noun was encountered; both the global context (the highly constraining birthday cake scenario) and local context (a semantic feature of cake) led the listener to anticipate the

word and its phonological form.³ When a mismatching word was encountered instead (VEGGIES), the globally predictable conditions (sweet/healthy and tasty CAKE) and the globally unpredictable/locally consistent (healthy and tasty VEGGIES) conditions showed increases in amplitude at right frontal electrode sites. These results suggest that listeners make specific lexical predictions during discourse processing and that the brain detects the mismatch between the predicted and actual inputs within 200 ms after stimulus onset.

The N200 effect may reflect the mismatch between the expected and received word forms. As we mentioned in the

³ To confirm that the globally and locally consistent condition (sweet and tasty CAKE) generated the strongest prediction for a particular word form, we conducted an additional offline rating task in which 40 participants completed the experimental stories, truncated at the critical nouns (i.e., after the feature words). The results showed that whereas over half of the participants (55 %) specifically predicted critical nouns like CAKE after the “sweet and tasty” condition, 0 % predicted critical nouns like VEGGIES in this condition. Therefore, “sweet and tasty VEGGIES” was the only condition in which a very strong prediction for a specific item was disconfirmed (and was the condition driving the N200 effect). It should be noted that although the cloze probability for the most predictable condition (sweet and tasty CAKE) was weaker in this norming test, when the feature words preceded the cutoff point, than when the feature words were not present (55 % vs. 78 %), 55 % still represented a specific prediction for over half of the participants on a given item, and more importantly, was significantly different from the next most predictable condition (55 % vs. 33 %, $p < .001$). In this condition, specific predictions for globally predictable nouns like CAKE were reduced after encountering features such as “healthy and tasty,” dropping the cloze probability for CAKE in this condition down (to 33 %). Therefore, critical nouns like VEGGIES in the latter condition were still unexpected (cloze probability of 1 %), but did not represent a mismatch with a specific prediction to the same extent as in the “sweet and tasty” condition.

Table 3 Event-related potential results for the critical nouns (omnibus analysis)

	df	N200			N400			PNP		
		F	p	η^2	F	p	η^2	F	p	η^2
Midline										
Global Predictability	(1, 18)	5.23	*	.06	48.02	***	.37	3.73	^	–
Local Consistency	(1, 18)	<1	–	–	6.75	*	.05	1.41	–	–
Global Predictability× Local Consistency	(1, 18)	2.85	–	–	3.97	^	–	5.44	*	.07
Global Predictability× Electrode	(3, 54)	2.02	–	–	2.1	–	–	<1	–	–
Local Consistency× Electrode	(3, 54)	<1	–	–	<1	–	–	1.22	–	–
Global Predictability× Local Consistency× Electrode	(3, 54)	3.43	^	^	<1	–	–	2.18	–	–
Medial										
Global Predictability	(1, 18)	2.69	–	–	40.14	***	.40	4.19	^	^
Local Consistency	(1, 18)	<1	–	–	7.07	*	.06	<1	–	–
Global Predictability× Local Consistency	(1, 18)	5.28	*	.06	5.18	*	.03	8.99	**	.11
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	<1	–	–
Local Consistency× Hemisphere	(1, 18)	<1	–	–	1.24	–	–	<1	–	–
Global Predictability× Local Consistency× Hemisphere	(1, 18)	<1	–	–	<1	–	–	1.8	–	–
Global Predictability× Anteriority	(2, 36)	1.66	–	–	4.33	*	.01	<1	–	–
Local Consistency× Anteriority	(2, 36)	<1	–	–	1.54	–	–	1.34	–	–
Global Predictability× Local Consistency× Anteriority	(2, 36)	1.34	–	–	<1	–	–	<1	–	–
Global Predictability× Hemisphere× Anteriority	(2, 36)	<1	–	–	<1	–	–	<1	–	–
Local Consistency× Hemisphere× Anteriority	(2, 36)	<1	–	–	1.27	–	–	3.59	*	.01
Global Predictability× Local Consistency× Hemisphere× Anteriority	(2, 36)	1.06	–	–	<1	–	–	1.35	–	–
Lateral										
Global Predictability	(1, 18)	4.64	*	.05	40.44	***	.35	3.29	–	–
Local Consistency	(1, 18)	<1	–	–	6.66	*	.05	<1	–	–
Global Predictability× Local Consistency	(1, 18)	4.66	*	.04	4.79	*	.02	9.97	**	.08
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	<1	–	–
Local Consistency× Hemisphere	(1, 18)	1.23	–	–	3.07	–	–	2.84	–	–
Global Predictability× Local Consistency× Hemisphere	(1, 18)	<1	–	–	<1	–	–	2.01	–	–
Global Predictability× Anteriority	(3, 54)	2.91	–	–	6.97	*	.01	<1	–	–
Local Consistency× Anteriority	(3, 54)	<1	–	–	1.84	–	–	4.81	*	.01
Global Predictability× Local Consistency× Anteriority	(3, 54)	1.63	–	–	1.82	–	–	<1	–	–
Global Predictability× Hemisphere× Anteriority	(3, 54)	3.33	*	.01	3.18	*	.00	2.61	–	–
Local Consistency× Hemisphere× Anteriority	(3, 54)	1.86	–	–	<1	–	–	<1	–	–
Global Predictability× Local Consistency× Hemisphere× Anteriority	(3, 54)	2.59	–	–	12.49	**		1.59	–	–

*** $p < .001$, ** $p < .01$, * $p < .05$, ^ $p < .07$

introduction, N200 effects with similar timings and topographic distributions have been reported in previous studies in which the incoming phonological input mismatched the form of the most expected word (Connolly & Phillips, 1994; van den Brink et al., 2001; see Diaz & Swaab, 2007, for a similar effect for words presented in lists rather than sentences). These N200 effects have several similarities with ERP effects that have been observed to mismatches between expected and reviewed stimuli in domains other than language processing.

For example, early frontal negativities have been reported in response to auditory “oddballs” in perceptual tasks

(Näätänen, 1995; Näätänen, Gaillard, & Mäntysalo, 1978) and to a wide range of novelty/expectedness manipulations in cognitive-control paradigms (see Folstein & Van Petten, 2008, for a recent review). Similarly, the “oddball N200” (elicited by infrequent stimuli), feedback ERN (elicited by similar manipulations involving favorable/unfavorable feedback), and response ERN (elicited by error and response conflict) all have similar frontal/frontal-central distributions across the scalp and timings (Holroyd, 2004; Holroyd & Coles, 2002; Yeung, Botvinick, & Cohen, 2004). As Folstein and Van Petten pointed out, the anterior N200 that is often observed in response to infrequent auditory stimuli

Table 4 Event-related potential results for the critical nouns: Simple effects of local consistency

	<i>df</i>	N200			N400			PNP		
		<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2
Globally Predictable: Locally Consistent vs. Locally Inconsistent										
Midline										
Local Consistency	(1, 18)	<1	–	–	<1	–	–	8.15	*	.22
Local Consistency × Electrode	(3, 54)	1.21	–	–	<1	–	–	<1	–	–
Medial										
Local Consistency	(1, 18)	<1	–	–	<1	–	–	9.56	**	.28
Local Consistency × Hemisphere	(1, 18)	<1	–	–	<1	–	–	<1	–	–
Local Consistency × Anteriority	(2, 36)	<1	–	–	<1	–	–	1.14	–	–
Local Consistency × Hemisphere × Anteriority	(2, 36)	<1	–	–	1.12	–	–	<1	–	–
Lateral										
Local Consistency	(1, 18)	1.38	–	–	<1	–	–	9.99	**	.2
Local Consistency × Hemisphere	(1, 18)	<1	–	–	3.15	–	–	<1	–	–
Local Consistency × Anteriority	(3, 54)	1.2	–	–	<1	–	–	1.69	–	–
Local Consistency × Hemisphere × Anteriority	(3, 54)	<1	–	–	3.44	*	.01	1.31	–	–
Globally Unpredictable: Locally Consistent vs. Locally Inconsistent										
Midline										
Local Consistency	(1, 18)	2.13	–	–	8.94	**	.26	<1	–	–
Local Consistency × Electrode	(3, 54)	2.52	–	–	<1	–	–	3.06	–	–
Medial										
Local Consistency	(1, 18)	4.29	^	–	11.32	**	.33	1.11	–	–
Local Consistency × Hemisphere	(1, 18)	2.23	–	–	1.17	–	–	1.38	–	–
Local Consistency × Anteriority	(2, 36)	<1	–	–	1.29	–	–	1.15	–	–
Local Consistency × Hemisphere × Anteriority	(2, 36)	1.11	–	–	<1	–	–	2.95	^	–
Lateral										
Local Consistency	(1, 18)	3.33	–	–	10.7	**	.25	1.08	–	–
Local Consistency × Hemisphere	(1, 18)	1.23	–	–	<1	–	–	5.02	*	.01
Local Consistency × Anteriority	(3, 54)	<1	–	–	3.17	^	–	3.81	^	–
Local Consistency × Hemisphere × Anteriority	(3, 54)	4.73	**	.01	5.16	**	.01	<1	–	–

*** $p < .001$, ** $p < .01$, * $p < .05$, ^ $p < .07$

may represent a mixture of several related components (MMN, N2b, N2c). Indeed, within the cognitive-control, anterior-negativity literature, there has been some debate as to whether or not these represent distinct effects and/or processes (Holroyd, 2004; Yeung et al., 2004). Across domains, many of the early, anterior negative deflections that have been observed in response to novel, unexpected, or error-related stimuli have been interpreted as reflecting some type of mismatch between expected and received inputs (Folstein & Van Petten, 2008; Holroyd, 2004; Holroyd & Coles, 2002; Yeung et al., 2004).

In the present study, we suggest that the increased early frontal negativity seen in the globally predictable/locally inconsistent condition (sweet and healthy VEGGIES) relative to the other conditions reflects an early cost associated with detecting a mismatch between the expected and received word forms. This pattern of results demonstrates that the activation

of a specific lexical form when the context is maximally constraining (both global and local cues point to the same word) leads to the detection of a mismatch (representational conflict) when a different word than anticipated is heard.

In summary, the results for the critical nouns in the N200 window show that listeners activate the specific phonological form when they have a strong basis on which to expect specific lexical items (as is the case for the “sweet and tasty VEGGIES” condition). When the incoming phonological input does not match the activated representation, an early frontal-central negativity is generated. This effect is (1) similar in timing and topography to other mismatch effects that have been observed, in both the language literature and the cognitive-control literature, and (2) topographically distinct from the N400 effects that follow.

Our N400 results showed graded effects of global predictability and local consistency, with the smallest N400 to

Table 5 Event-related potential results for the critical nouns: Simple effects of global predictability

	<i>df</i>	N200			N400			PNP		
		<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2
Locally Consistent: Globally Predictable vs. Globally Unpredictable										
Midline										
Global Predictability	(1, 18)	<1	–	–	19.72	***	.34	<1	–	–
Global Predictability× Electrode	(3, 54)	3.8	*	.07	1.12	–	–	<1	–	–
Medial										
Global Predictability	(1, 18)	<1	–	–	17.54	***	.4	1.06	–	–
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	1.23	–	–
Global Predictability× Anteriority	(2, 36)	2.54	–	–	2.84	–	–	1.08	–	–
Global Predictability× Hemisphere× Anteriority	(2, 36)	1.82	–	–	<1	–	–	<1	–	–
Lateral										
Global Predictability	(1, 18)	<1	–	–	17.55	***	.32	1.33	–	–
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	<1	–	–
Global Predictability× Anteriority	(3, 54)	2.93	–	–	1.66	–	–	<1	–	–
Global Predictability× Hemisphere× Anteriority	(3, 54)	1.88	–	–	2.79	^	–	2.91	^	–
Locally Inconsistent: Globally Predictable vs. Globally Unpredictable										
Midline										
Global Predictability	(1, 18)	6.23	*	.19	31.65	***	.56	7.36	*	.23
Global Predictability× Electrode	(3, 54)	<1	–	–	1.54	–	–	2.55	–	–
Medial										
Global Predictability	(1, 18)	6.48	*	.23	33.74	***	.61	4.76	*	.16
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	<1	–	–
Global Predictability× Anteriority	(2, 36)	<1	–	–	3.27	–	–	1.44	–	–
Global Predictability× Hemisphere× Anteriority	(2, 36)	<1	–	–	<1	–	–	1.36	–	–
Lateral										
Global Predictability	(1, 18)	9.79	**	.24	34.83	***	.54	11.54	**	.23
Global Predictability× Hemisphere	(1, 18)	<1	–	–	<1	–	–	1.01	–	–
Global Predictability× Anteriority	(3, 54)	1.91	–	–	9.16	**	.02	1.74	–	–
Global Predictability× Hemisphere× Anteriority	(3, 54)	3.87	*	.01	13.62	***	.01	1.11	–	–

*** $p < .001$, ** $p < .01$, * $p < .05$, ^ $p < .07$

globally predictable, locally consistent nouns (sweet and tasty CAKE), followed by globally predictable, locally inconsistent nouns (healthy and tasty CAKE), then by globally unpredictable, locally consistent nouns (healthy and tasty VEGGIES), and finally by globally unpredictable, locally inconsistent nouns (sweet and tasty VEGGIES).

The globally predictable, locally consistent (sweet and tasty CAKE) condition displayed a floor effect in this time window. This is the only condition in which a specific lexical prediction was fully supported: The critical noun was highly predictable at the global level, and was further supported by the feature word preceding the noun. If listeners use context to anticipate specific lexical items, they are best equipped to do so in the maximally constraining condition, in which enough context is present to narrow down expectations to a particular form. We do not suggest that the prediction of lexical forms is common in language processing. Rather, as several recent

accounts have suggested, it is likely that readers/listeners activate the general semantic features (e.g., animacy) of upcoming input unless the context is constraining enough to license a more specific prediction (Kuperberg, 2013; Pickering & Garrod, 2013; Szewczyk & Schriefers, 2013; Van Petten & Luka, 2012). We suggest that the globally predictable/locally consistent condition in the present experiment is just such a case, leading to little if any access/retrieval processing when the actual critical word is heard. This is reflected by the amplitude of the N400 being maximally attenuated relative to all other conditions.

In contrast, the N400 to the globally predictable, locally inconsistent (healthy and tasty CAKE) condition was slightly, but significantly, larger, although still substantially attenuated relative to the two globally unpredictable conditions (sweet/healthy and tasty VEGGIES). This indicates that the presence of the inconsistent feature words (“healthy” before CAKE)

had a significant influence on processing, such that the critical nouns (CAKE) that followed did not enjoy the same level of facilitation as the globally predictable words that had followed consistent feature words. Likewise, the presence of a locally consistent feature significantly reduced the N400 to globally unpredictable words (healthy and tasty VEGGIES) relative to when unpredictable words were preceded by locally inconsistent feature words (sweet and tasty VEGGIES). Thus, global and local context both contributed to the degree of match/mismatch between the expected and received semantic content (features) at the critical nouns, as reflected by the graded pattern of N400 results. This is consistent with previous work in which global and local context were manipulated within the same paradigm (Boudewyn et al., 2012; Boudewyn et al., 2013; Camblin et al., 2007). In those studies, listeners were rapidly sensitive to both the discourse message and the presence of local primes that preceded the target words. The present results expand on these findings to show that, even in constraining discourse contexts in which a specific word is highly predictable (CAKE), listeners dynamically adjust their expectations after encountering locally inconsistent input (healthy). Overall, the results in the N400 window at the critical nouns show (1) evidence for the activation of globally predictable words and (2) evidence that both global and local context influence processing.

Conflict and adaptation

Global predictability and local consistency interacted significantly in the PNP time window. Local consistency affected the processing of globally predictable nouns, such that globally predictable, locally inconsistent nouns (healthy and tasty CAKE) showed a larger PNP than did locally consistent nouns (sweet and tasty CAKE). Likewise, local consistency affected the processing of globally unpredictable nouns: Globally unpredictable, locally consistent nouns (healthy and tasty VEGGIES) showed a larger PNP than did locally inconsistent nouns (sweet and tasty VEGGIES). No significant PNP effect of global predictability emerged for locally consistent nouns (sweet and tasty CAKE vs. healthy and tasty VEGGIES); this was the only comparison in which no PNP effect was found. However, global predictability affected the processing of locally inconsistent nouns, such that globally predictable, locally inconsistent nouns (healthy and tasty CAKE) showed a larger PNP than did globally unpredictable, locally inconsistent nouns (sweet and tasty VEGGIES).

We refer to the late positive effects in this study as *post-N400 positivities* (PNPs), because this is a theory-neutral term that does not exclude the possibility that multiple ERP effects may occur in this time window (see Van Petten & Luka, 2012). The PNP effects were quite late (900–1,200 ms post-word-onset) in comparison to the typical latency of either the P600 effect or the PNP that has recently been linked to the

processing of unpredictable input, both of which are commonly observed in the range of 500–900 ms post-word-onset⁴ (see Van Petten & Luka, 2012, for a recent review). However, as will be discussed in more detail below, the present PNP effects share a number of similarities with other late positive effects that have been observed in language-processing paradigms, particularly the P600 effect and the frontal PNP that has been linked to the processing of unpredictable input. In addition to the positive polarity and the relative timing (post-N400), the present PNP effects appear to be driven by increased demands on revision, updating, or conflict-monitoring/resolution processes, all of which have been previously related to P600/PNP effects (e.g., Brothers et al., 2015; Federmeier et al., 2007; Friederici, 2002; Kolk & Chwilla, 2007; Kuperberg, 2007; O'Rourke & Van Petten, 2011; Thornhill & Van Petten, 2012).

The P600 is increased in response to syntactic errors (e.g., Osterhout & Mobley, 1995), syntactic complexity (e.g., Kaan, Harris, Gibson, & Holcomb, 2000), and conflict at the syntax–semantics interface (e.g., Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). It is typically maximal over central-posterior/parietal electrode sites. There is some debate over the precise functional significance of the P600. One view is that it reflects revision, or attempts at revision, of previously adopted syntactic structures (Friederici, 2002; Hahne & Friederici, 1999). Another view attributes a domain-general error-monitoring and reprocessing function to the P600 (Kolk & Chwilla, 2007; van Herten, Chwilla, & Kolk, 2006). Following the finding of “semantic” P600 effects in response to words that appeared in simple, syntactically well-formed sentences, but that represented thematic role violations (e.g., “At breakfast the eggs would eat”), it has been suggested that the P600 may reflect continued combinatorial processing following a conflict between semantic and syntactic processing streams (Kuperberg, 2007; Kuperberg et al., 2003). The latter account has since been refined to posit that the P600 represents the processing costs of disconfirmed predictions about events or event structures (Kuperberg, 2013). Specifically, Kuperberg (2013) suggested that a processing cost is incurred when a specific “high-certainty” prediction is disconfirmed (whether that prediction concerns an event, structure, or thematic role assignment), and that this cost is reflected in the P600.

In contrast, PNPs with a more frontal distribution across the scalp have been observed in response to plausible words that are unpredictable in context (see Thornhill & Van Petten, 2012; Van Petten & Luka, 2012, for overviews). Unlike the central-posterior/parietal P600 effect, which may or may not occur after an N400 effect, the frontal PNPs follow N400 effects. This underscores the PNP's link to the processing of

⁴ However, late PNP effects have been observed extending out to our rather late time range in several studies (Otten & Van Berkum, 2008; Thornhill & Van Petten, 2012).

input that is semantically unpredictable. Although recent retrospective reviews of the N400 literature have revealed that these frontal PNPs have appeared in a number of studies, they have not received much attention until recently (DeLong, Urbach, Groppe, & Kutas, 2011; Federmeier et al., 2007; but see Kutas, 1993; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012). As such, the functional significance of the frontal PNP is open to interpretation. One possibility is that it reflects processing costs associated with encountering input that was not predicted but is plausible in context (Van Petten & Luka, 2012). As to the specific nature of the processing costs, possibilities include (1) inhibition of the predicted word (since a different word was encountered in its place; Kutas, 1993) and (2) additional processing needed to discard the previous predictions (which were disconfirmed by plausible, but unpredictable, input) and to update representations of the context accordingly (Kuperberg, 2013).

In the present study, two of the effects in the PNP window had a parietal distribution that more closely resembled the P600 than the frontal PNP: the effect of local consistency for globally predictable words (sweet and tasty CAKE vs. healthy and tasty CAKE) and the effect of global predictability for locally inconsistent words (sweet and tasty VEGGIES vs. healthy and tasty CAKE). The third PNP effect was smaller and had a more frontal distribution that more closely resembled the frontal PNP than the P600: the effect of local consistency for globally unpredictable words (sweet and tasty VEGGIES vs. healthy and tasty VEGGIES). Despite small distributional differences, a parsimonious explanation of these effects is that they reflect context-updating processes, triggered by conflict between expectations that were based on global context and those based on the most recent, local context. When the global and local contexts were consistent, no late PNP was observed. Importantly, this was true for the doubly supported critical nouns (sweet and tasty CAKE) and for the nouns that were doubly unsupported (sweet and tasty VEGGIES). As can be seen in Fig. 2, the waveforms for these two conditions overlap in the PNP time window, despite having diverged in the preceding N200 and N400 time windows and despite being at the opposite extremes of the predictability manipulation. This may be because neither of these conditions triggered additional context-updating processes, which as Kuperberg (2013) suggested may involve discarding previous predictions in the face of alternative input that is plausible enough to integrate. In the case of the doubly supported condition, there was no context-updating cost, because the input was consistent with both scenario-based and lexically based expectations. In the case of the doubly unsupported condition, there was no context to support abandoning the developing scenario; the given phrase was not a plausible continuation to the story. Thus, it did not provide enough information to update the developing representation. In other words, the phrase

may have been interpreted as an anomaly in a birthday cake scenario and did not trigger an updating process in which the scenario was reinterpreted.

The support for context/expectation updating was present in the “single-conflict” conditions. In these conditions (globally predictable/locally inconsistent [healthy and tasty CAKE] and globally unpredictable/locally consistent [healthy and tasty VEGGIES]), a local cue was present that, when combined with the noun, warranted updating the discourse representation. In the case of the globally supported, locally inconsistent condition (healthy and tasty CAKE), the presence of the inconsistent feature continuation in an otherwise supportive global context triggered efforts to integrate the incompatible feature into the representation of the discourse (to accommodate what was an atypical exemplar of the noun). The PNP in this condition showed up in two of the comparisons: It was significantly larger than in both the globally supported/locally consistent condition (sweet and tasty CAKE) and the globally unsupported/locally inconsistent condition (sweet and tasty VEGGIES). In fact, as can be seen in Fig. 2, the PNP in this condition was the largest and most broadly distributed PNP that was observed in the study.

In the globally unsupported, locally consistent condition (healthy and tasty VEGGIES), the presence of the consistent feature may have provided enough support that an attempt at updating was made, even though the noun that followed a few words later did not fit well with the global context. This possibility is consistent with the idea that plausible, but unpredictable, input may trigger listeners/readers to abandon their expectations and update their representations accordingly. In this condition, the local cue appears to have lent some plausibility to the noun, which would otherwise be a complete mismatch to the preceding context. Responses in this condition significantly diverged from the doubly unsupported condition (sweet and tasty VEGGIES) in the PNP window, consistent with the idea that the PNP is specific to circumstances in which some contextual support is available to trigger updating processes. However, this condition only marginally diverged from the doubly supported condition (sweet and tasty CAKE). Although some attempts at updating and integration might have been possible, the globally unsupported, locally supported (healthy and tasty VEGGIES) noun still did not fit the overall scenario (birthday party) very well. It may have been more feasible to reconcile the most predictable noun (CAKE) with an atypical semantic feature (healthy) than to accommodate a noun that was a poor fit to the global context (VEGGIES), even with local support (healthy). If so, this would explain the more robust PNP that was observed in the former than in the latter condition, and why the latter condition showed only a marginal PNP relative to the fully supported condition.

Conclusions

Our results support four main conclusions. First, the N400 effect that was observed at the feature words is consistent with recent suggestions that access/retrieval is facilitated when some or all relevant semantic features have been activated by context (Kuperberg, 2013). Second, the graded N400 effect to critical nouns shows that expectations for upcoming words can be modulated by constraints in the global discourse and by the meanings of words in the local context, even when the discourse context is highly predictive of specific lexical items. Furthermore, the results at the critical nouns show that listeners dynamically adjust their expectations as incoming words are heard, since the feature words influenced processing of the critical nouns that appeared downstream, strengthening the expectation for the globally predictable nouns when they were consistent, and weakening it when they were inconsistent. Third, when the context is sufficiently constraining, as when both scenario-based expectations and the local context converge, a specific lexical item (including semantic features and word form) is activated. Evidence for this comes from the N200 effect seen at the nouns, in which a mismatch between the predicted and perceived lexical forms resulted in an early, frontal negative deflection (distinct from the N400). Fourth, listeners dynamically adjust their expectations when input begins to diverge from the representation of the context that has been developing up to that point. Critically, the inconsistent input that triggers the adjustment and updating processes must have some level of support from prior context: Input that is completely inconsistent/implausible (i.e., at both the global and local levels) does not appear to elicit revision/updating processes. Evidence for this comes from the PNP effects that were observed at the critical nouns, in which a late positive shift was elicited only in those conditions in which the feature words were inconsistent with the most predictable completion. We suggest that the conflict triggered updating processes at the subsequently encountered noun. This suggestion is consistent with models of discourse processing in which coherence breaks are posited to trigger “shifting” processes (Gernsbacher, 1996, 1997). Inconsistencies or changes in coherence can serve as useful information to the comprehender, but are associated with more effortful processing. The adjustment of expectations is critical during comprehension, because incoming words do not always map directly to what is expected on the basis of prior context. Discourse representations are not static, but are continually updated as new information is encountered, and therefore the processing of incoming words is dynamically facilitated by the most current contextual representation that is available.

References

- Boudewyn, M. A., Gordon, P. C., Long, D., Polse, L., & Swaab, T. Y. (2012). Does discourse congruence influence spoken language comprehension before lexical association? Evidence from event-related potentials. *Language and Cognitive Processes, 27*, 698–733. doi:10.1080/01690965.2011.577980
- Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology, 4*, 60. doi:10.3389/fpsyg.2013.00060
- 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.
- Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*, 103–128. doi:10.1016/j.jml.2006.07.005
- Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6*, 256–266.
- DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology, 48*, 1203–1207.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117–1121. doi:10.1038/nn1504
- Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain Research, 1146*, 85–100. doi:10.1016/j.brainres.2006.07.034
- Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505. doi:10.1111/j.1469-8986.2007.00531.x
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495. doi:10.1006/jmla.1999.2660
- 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*, 149–161. doi:10.1016/j.bandl.2010.07.006
- 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. doi:10.1016/j.brainres.2006.06.101
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45*, 152–170.
- Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465–495. doi:10.1080/14640748108400804
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78–84.
- Gernsbacher, M. A. (1996). The structure-building framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 289–311). Mahwah, NJ: Erlbaum.
- Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23*, 265–304.
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.

- Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. *Neurophysiology*, 78, 447–455.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709. doi:10.1037/0033-295X.109.4.679
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York, NY: Oxford University Press.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201.
- Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, 100, 257–261. doi:10.1016/j.bandl.2006.07.006
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49. doi:10.1016/j.brainres.2006.12.063
- Kuperberg, G. R. (2013). The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 176–192). Baltimore, MD: Paul Brookes.
- Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17, 117–129. doi:10.1016/S0926-6410(03)00086-7
- Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, 8, 533–572.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4, 463–470.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. doi:10.1146/annurev.psych.093008.131123
- Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, 25, 484–502. doi:10.1162/jocn_a_00328
- Lau, E. F., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98, 74–88. doi:10.1016/j.bandl.2006.02.003
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213. doi:10.3389/fnhum.2014.00213
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, 62, 203–208.
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, 37, 547–559. doi:10.3758/BF03192726
- Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, 66, 545–567.
- Näätänen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing*, 16, 6–18.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313–329.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, 24, 691–701. doi:10.1016/j.cogbrainres.2005.04.003
- O'Rourke, P. L., & Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain Research*, 1392, 62–79.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Otten, M., & Van Berkum, J. J. A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research*, 1153, 166–177. doi:10.1016/j.brainres.2007.03.058
- Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: prediction or priming? *Discourse Processes*, 45, 464–496. doi:10.1080/01638530802356463
- Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, 67, 426–448. doi:10.1016/j.jml.2012.07.003
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–347. doi:10.1017/S0140525X12001495
- Ruchkin, D. S., Johnson, R., & Friedman, D. (1999). Scaling is necessary when making comparisons between shapes of event-related potential topographies: A reply to Haig et al. *Psychophysiology*, 36, 832–834.
- Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 397–440). New York, NY: Oxford University Press.
- Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language*, 68, 297–314. doi:10.1016/j.jml.2012.12.002
- 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, 382–392.
- Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology*, 39, 791–808.
- 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, and Cognition*, 31, 443–467. doi:10.1037/0278-7393.31.3.443
- van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11, 657–671. doi:10.1162/08992999563724
- van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, 17, 701–718. doi:10.1016/S0926-6410(03)00196-4
- van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, 13, 967–985.
- van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, 18, 1181–1197. doi:10.1162/jocn.2006.18.7.1181

- Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8*, 485–531.
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190. doi:10.1016/j.ijpsycho.2011.09.015
- Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*, 238–242.
- 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*, 1272–1288. doi:10.1162/0898929041920487
- Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111*, 931–959. doi:10.1037/0033-295X.111.4.931