Predicting form and meaning: Evidence from brain potentials

Aine Ito *, Martin Corley, Martin J. Pickering, Andrea E. Martin, Mante S. Nieuwland

Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom

A R T I C L E   I N F O

A B S T R A C T

We used ERPs to investigate the pre-activation of form and meaning in language comprehension. Participants read high-cloze sentence contexts (e.g., "*The student is going to the library to borrow a…*"), followed by a word that was predictable (*book*), form-related (*hook*) or semantically related (*page*) to the predictable word, or unrelated (*sofa*). At a 500 ms SOA (Experiment 1), semantically related words, but not form-related words, elicited a reduced N400 compared to unrelated words. At a 700 ms SOA (Experiment 2), semantically related words and form-related words elicited reduced N400 effects, but the effect for form-related words occurred in very high-cloze sentences only. At both SOAs, form-related words elicited an enhanced, post-N400 posterior positivity (Late Positive Component effect). The N400 effects suggest that readers can pre-activate meaning *and* form information for highly predictable words, but form pre-activation is more limited than meaning pre-activation. The post-N400 LPC effect suggests that participants detected the form similarity between expected and encountered input. Pre-activation of word forms crucially depends upon the time that readers have to make predictions, in line with production-based accounts of linguistic prediction.

© 2015 Elsevier Inc. All rights reserved.

Introduction

People regularly use contextual information and world knowledge to predict aspects of language that are likely to be mentioned as a sentence or discourse unfolds (e.g., Altmann & Kamide, 1999; Federmeier, 2007; Huettig, 2015; Kutas, DeLong, & Smith, 2011). Prediction is often hypothesized to occur via a so-called *pre-activation* mechanism, whereby some aspects of word meaning, grammar or form are activated before the onset of the predicted word (e.g., DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; Laszlo & Federmeier, 2009; Otten, Nieuwland, & Van Berkum, 2007; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). But how these types of linguistic information are pre-activated is still

unclear. The production-based prediction account proposes prediction via a comprehender's production system (Pickering & Garrod, 2007, 2013). Under this account, pre-activation of form does not occur in the absence of pre-activation of meaning, because the language production system first accesses meaning, and then maps the meaning information onto form information. In this paper, we report two event-related brain potential (ERP) experiments that investigate pre-activation of meaning and form of predictable words during language comprehension to explore the relationship between meaning and form pre-activation. We investigate pre-activation, as indexed by N400 ERP modulations (Kutas & Federmeier, 2011), at a word presentation rate that is standard in reading ERP studies (Experiment 1; 500 ms per word) and at a slower presentation rate (Experiment 2; 700 ms per word) which allows more time to generate online predictions. Below, we first discuss the production-based prediction theory, and then outline existing evidence for the pre-activation of meaning and of form before introducing the current study.

---

* Corresponding author at: Department of Psychology, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland, United Kingdom.

E-mail address: aine.ito@ed.ac.uk (A. Ito).

Production-based prediction accounts

Pickering and Garrod (2007, 2013) proposed that people use the language production system when predicting upcoming words during comprehension. According to this account, when people comprehend sentences, they covertly imitate those sentences and implement their production systems to predict upcoming words. Lexical prediction is thought to involve pre-activation of linguistic information (e.g., word form, meaning) of predictable words. Linguistic information associated with predictable words is pre-activated using the same mechanisms that are used to produce words.

A most parsimonious possibility is that comprehenders make direct use of the mechanisms involved in language production – a version of prediction-by-production that we call *prediction-with-implementation*. Although language production models (e.g., Dell & O'Seaghdha, 1992; Levelt, Roelofs, & Meyer, 1999) differ in many important respects, they agree on the view that people produce a word by first activating its semantic information and then proceeding through stages that lead to activation of its phonological or orthographic information (its word form). These stages take several hundred milliseconds according to most estimates (see Indefrey & Levelt, 2004). According to prediction-with-implementation, comprehenders also pre-activate semantic information before form information, following roughly the same time-course. It is of course possible for the comprehender to actually complete the speaker's utterance, simply by continuing the process of production until the stage of articulation – this is exactly what happens in a cloze test.

However, full implementation of the production system for prediction requires time and resources. When these are lacking, only a part of the production system may be used for prediction. As activation of form information follows activation of semantic information in the language production system, a partly engaged production system might lead to pre-activation of semantic information but not of form information. This means that a comprehender might pre-activate meaning without pre-activating form under conditions of difficulty, but would not pre-activate form without pre-activating meaning.[1]

However, we note that a pattern wherein meaning pre-activation is more likely to occur than form pre-activation could also be compatible with an alternative account involving cascaded pre-activation. Cascaded pre-activation

has not previously been hypothesized to underlie prediction, but cascaded activation is a common mechanism in theories of spoken word recognition (Marslen-Wilson, 1987; Norris, 1994) and in theories of language production (Caramazza, 1997; Dell, 1986). Pre-activation of meaning may cascade into pre-activation of word form, whether or not predictions are generated by the production system. The ramifications of this account will be further discussed in the General Discussion.

Predicting meaning

Classic findings from Kutas and colleagues have shown that anomalous words lead to increased N400 ERPs in comparison to plausible words in the same sentence contexts (Kutas & Hillyard, 1980) and that this N400 effect is reduced for words that are semantically related to the plausible word (Federmeier & Kutas, 1999; Kutas & Hillyard, 1984). Federmeier and Kutas (1999) had participants read discourse contexts that led them to expect a particular target word (e.g., "*They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of...*") and then presented them with that predictable word (*palms*), a related implausible word within the same semantic category as the target word (*pines*), or an implausible word from a different semantic category (*tulips*). As in Kutas and Hillyard (1984), participants' N400 responses were reduced for implausible within-category words (*pines*) compared to between-category words (*tulips*). This reduction was greater in *high-cloze* contexts, which were created based on a sentence completion pre-test in which participants produced the expected completion (*palms*) 90% of the time, than in *medium-cloze* contexts, in which participants produced the expected completion 59% of the time (see Thornhill & Van Petten, 2012, for similar findings).

Although implausible within-category words (*pines*) elicited a greater N400 reduction in high-cloze sentences than in medium-cloze sentences, they were rated as less plausible in high- than in medium-cloze sentences. Crucially then, because the N400 reduction did not pattern with the plausibility pre-test data, Federmeier and Kutas (1999) could rule out an integration account in which the observed N400 reductions reflected within-category words being more plausible sentence continuations (and therefore easier to integrate) than between-category words. They concluded that, prior to the onsets of the target words, participants had activated semantic features of the expected sentence continuations. This in turn implied activation of some of the within-category words' semantic features, resulting in facilitation of the within-category words relative to those which didn't share a semantic category, as indexed by N400 reduction.

Federmeier and Kutas' (1999) findings are indeed consistent with an account of prediction that operates via pre-activation of semantic category features. However, a remaining inconsistency comes from the fact that a pre-activation account also strongly suggests that high-cloze target words themselves should show a reduced N400 effect compared to medium-cloze target words. But, surprisingly, Federmeier and Kutas (1999) did not find this basic effect of cloze probability.

---

[1] Pickering and Garrod (2013) in fact proposed a different type of production-based prediction that they called *prediction-by-simulation*. To summarize briefly, there is good evidence that people predict their own utterances using so-called forward models, based on associations between their intention (e.g., to talk about a kite) and aspects of the word they would use to describe that intention (e.g., the phoneme /k/). These forward models are ready before the utterance itself (thus allowing self-monitoring), and there is no reason that predictions of meaning need be ready before predictions of form. They can then use such forward models to predict during comprehension, again before the speaker produces the utterance. This form of prediction makes no claim that prediction should depend on time or resources, and in particular does not assume that prediction of form is less likely to occur than prediction of meaning. However, Pickering and Garrod's model is compatible with the occurrence of both prediction-by-simulation and prediction-with-implementation.

A possible alternative explanation of Federmeier and Kutas' data is that the N400 reduction for implausible within-category words (*pines*) occurred because within-category words also receive substantial lexical priming from the individual words in a sentence context that is highly supportive of the target word (*palms*). If, in sentence contexts that are high-cloze for the target word, within-category words also have a strong semantic relationship with the words in the sentence context (compared to the same within-category words in low-cloze contexts), a larger N400 reduction for within-category words would be observed. But it would reflect facilitation of the within-category word through lexical priming[2] rather than through semantic pre-activation of the target word.

Importantly, Metusalem et al. (2012) found that semantic pre-activation is not limited to semantic category features of expected words, while ruling out an explanation in terms of lexical priming. They conducted a pre-test to establish words that were commonly associated with the discourses presented in their experiment. They reported an N400 reduction for anomalous words that were associated with the events described in the discourses, but not to the expected target word itself (e.g., *jacket* is related to the event of building a snowman in the winter, but not to the concept 'snowman' itself), relative to event-unrelated anomalous words (*towel*). Because they controlled for degree of semantic association, this N400 difference between event-related and event-unrelated anomalous words could not be attributed to lexical priming of event-related words by context words. Though Metusalem et al. (2012) only used high-cloze sentences, and did not examine effects of cloze probability, their findings suggest that pre-activation of general or event-based knowledge relevant to the described event forms a basis of prediction (see also Nieuwland, 2015). In conclusion, people do appear to pre-activate semantic features of highly predictable upcoming words and semantic information that is more broadly relevant to the discourse context.

Predicting form

In contrast to pre-activation of meaning, it seems that pre-activation of form (i.e., what upcoming words will sound or look like) requires the prediction of a specific lexical item (as context words would not usually be related in form to a predictable word). A lexical prediction might pre-activate particular form features, which could in turn facilitate the processing of form-related words. In this paper, we do not distinguish prediction of sound (phonological form) and shape (orthographic form).

The evidence for form pre-activation is quite complex. Evidence for form-related N400 reduction comes from DeLong et al. (2005), who took advantage of the English phonotactic rule that the article *a* precedes consonant-initial words and the article *an* precedes vowel-initial

words, to investigate whether people pre-activate aspects of word form. Participants read high-cloze sentence contexts (e.g., "*The day was breezy so the boy went outside to fly*…") followed by the predictable article-noun combination (*a* followed by *kite*) or an unpredictable but plausible article-noun combination (*an* followed by *airplane*). The predictable noun began with a vowel and the unpredictable noun with a consonant, or vice versa. As expected, the N400 relative to the noun onset was reduced when the noun was predictable. Critically, the same effect was found at the preceding article when it was predictable (and thus, matched the noun). DeLong et al. argued that participants pre-activated form representations (e.g., an initial consonant) of the upcoming noun before the appearance of the noun (for similar results, see DeLong, Groppe, Urbach, & Kutas, 2012). These results therefore suggest that word form (at least the first phoneme of the word, but crucially not merely aspects of meaning) is pre-activated.

However, another possibility is that participants predicted the articles themselves, rather than predicting the noun and using the phonology of the noun to compute the article.[3] Recent evidence indicates that frequently occurring word sequences are comprehended more quickly than would be expected on the basis of their individual frequencies (Arnon & Snider, 2010; Tremblay, Derwing, Libben, & Westbury, 2011), suggesting that common sequences are represented (alongside individual words), in the mental lexicon. People may thus store article-noun sequences and use context to predict such sequences, rather than word form.

Laszlo and Federmeier (2009) had participants read predictable contexts (e.g., "*Before lunch he has to deposit his paycheck at the*…") that were completed by a predictable word (*bank*), words that were orthographically related to the predictable word (*bark*), pseudowords that were orthographically related to the predictable word (*pank*), and illegal strings that were orthographically related to the predictable word (*bxnk*). In matched unrelated conditions, participants read other predictable contexts (e.g., "*She loves the way the leaves change colors in the*…") that were completed by the predictable word (*fall*), and orthographically unrelated sets of words (*hook*), pseudowords (*jank*), and illegal strings (*tknt*). For words, pseudowords, and illegal strings, Laszlo and Federmeier found a reduced N400 for the forms that were orthographically related to the predictable word as compared to the forms that were orthographically unrelated. They concluded that pre-activated orthographical features impact semantic processing prior to any filter on lexical status.

In a related study, Kim and Lai (2012) found no N400 effect for pseudowords that were orthographically similar to predictable words (e.g., "*She measured the flour so she could bake a ceke*"), whereas dissimilar pseudowords (*tont*) elicited a clear N400 effect. Orthographically similar pseudowords elicited an enhanced P600, which was smaller than the P600 elicited by illegal strings (*srdt*). Similarly to Laszlo and Federmeier (2009), Kim and Lai argued that

---

[2] Within-category words were never lexical associates of the target words, but association norms include only strongly semantically related items. Importantly, lexical priming can also occur from semantically related or lexically co-occurring words which are non-associated (e.g., Hare, Elman, Tabaczynski, & McRae, 2009).

[3] Importantly, the N400 effect for unexpected articles in DeLong et al. (2005) was correlated with the predictability of the article, rather than the predictability of the subsequent noun. Hence, effects at the articles can be observed regardless of the cloze value of the noun.

the impact of prediction occurs before visual word recognition (because pseudowords cannot be recognized as words).

Laszlo and Federmeier (2009) and Kim and Lai (2012) interpreted the respective reduced or absent N400 effects associated with orthographical overlap as evidence for pre-activation of orthographic information. However, both studies used high proportions of pseudowords and non-words (54% in Laszlo & Federmeier, 2009; 75% in Kim & Lai, 2012). It remains unknown whether the reported effects generalize to settings involving only lexical items. Second, participants might learn to predict the occurrence of a nonword or a real word, depending entirely on the proportion of each type of stimulus and the nature of the design. What stimuli participants may learn to track can then affect the component that is elicited (Holcomb, 1988). This concern is particularly important for experiments concerned with prediction.

Another important concern is that these studies required participants to make a judgement or perform a task in addition to natural reading comprehension. For example, participants in Laszlo and Federmeier (2009) judged whether a stimulus was a "normal English sentence," and most non-predictable conditions were correctly answered with a "no" response (75% of the responses). Critical words can elicit positive deflections, such as the P300, as a function of extended task-related processing of these words (e.g., Newman, Connolly, Service, & Mcivor, 2003). These components could obscure N400 modulations via component overlap due to summation of positive and negative potentials at the scalp. Given that whether words were orthographically related or not was task-relevant (task difficulty increases if a critical word looks like the target word), the reported effects might reflect differences in task-related ERPs rather than, or in addition to, N400 differences.

Studies using non-words that are closely related in form to predictable words have also revealed post-N400 positive ERP effects (Late Positive Component or LPC effects) that seem to indicate that comprehenders consider the form of the predictable word. Newman and Connolly (2004) and Vissers, Chwilla, and Kolk (2006) found that pseudohomophones that were orthographically similar to highly predictable words (e.g., *bouks* for the predictable *books*) elicited larger LPCs than predictable words, but pseudohomophones that were orthographically similar to unpredictable words did not. Similarly, Laszlo and Federmeier (2009) and Kim and Lai (2012) reported a post-N400 LPC effect for pseudowords that were orthographically similar to the predicted words. Along with Vissers et al. (2006), both sets of authors interpreted the effect as a detection of a conflict between predicted and actually encountered words.

It is not yet clear whether the N400 and LPC effects previously interpreted as being due to form overlap would occur in the absence of a task that requires explicit evaluation of critical words, using a design with only real words.

The current study

We examined pre-activation of form and meaning as participants read for comprehension. To examine the effects of prediction, we assessed the N400 effects for high-cloze items and medium-cloze items (cf. Federmeier & Kutas, 1999). Participants read constraining sentences with the predictable word (*predictable condition*), an anomalous word sharing form features (i.e., having phonological/orthographic overlap) with the predictable word (*form condition*), an anomalous word that was semantically related to the predictable word (*semantic condition*), or an anomalous unrelated word (*unrelated condition*; see Fig. 1). The current study minimized potential artefactual effects by using real words only and employing no task related to critical words, while controlling for relevant variables. Moreover, we controlled the form-similarity of the semantically related words and the semantic relatedness of form-related words, to show that any demonstration of pre-activation of form cannot be wrongly ascribed to pre-activation of meaning, and vice versa.

We investigated whether there were N400 reductions for semantically related words and for form-related words, relative to the unrelated baseline. We expected the N400 reduction for both types of related words to be larger when predictable words were more strongly predicted (high cloze) than less strongly predicted (low cloze). We hypothesized that even if pre-activation of form features was weak or absent, form-related words might impact comprehension if people detect the conflict between actual input and predictable words. If so, we expected to find a post-N400 LPC effect, which should be strongest in highest cloze sentences, because the conflict should be greater when expectation for a specific word is stronger.

We conducted two experiments that differed in presentation rate. In Experiment 1, the stimulus onset asynchrony (SOA) was 500 ms; in Experiment 2, it was increased to 700 ms. Assuming that comprehenders make predictions by going through some of the stages that they use to produce utterances, then they might be unable to predict both meaning and form in Experiment 1, in part because of the relatively short time-lag and in part because comprehension would be rendered difficult by having to integrate all the words in the prior context. In contrast, we hypothesized that they would be able to predict both meaning and form in Experiment 2, given the longer SOA.
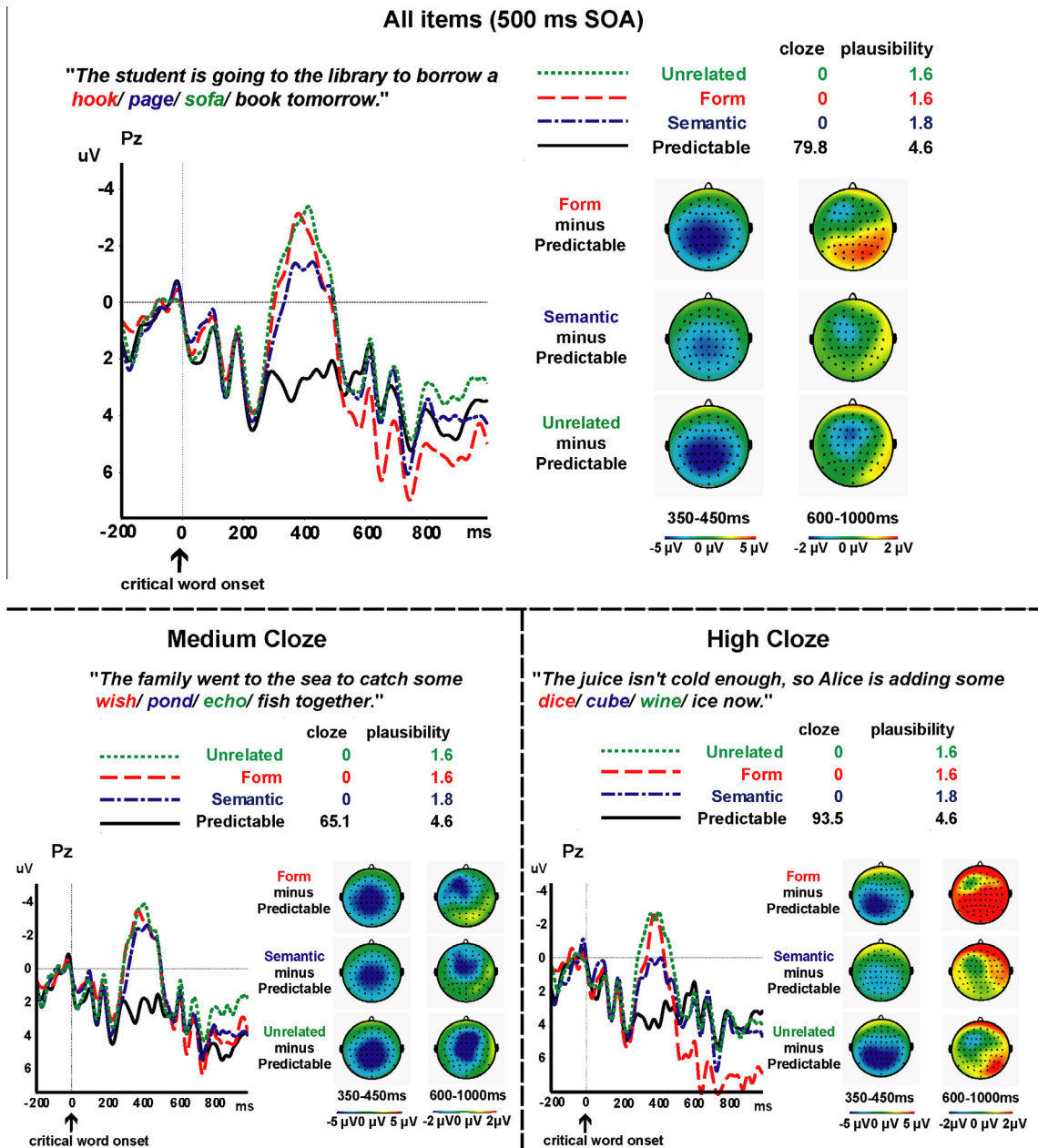
Experiment 1

Methods

Participants

Twenty-four English monolinguals (6 males and 18 females, age $M = 21.4$ years, $SD = 2.8$) took part in the experiment, having given informed consent. All participants were right-handed and free from neurological or language disorders.

Stimuli and experimental design

We constructed 160 items (from a candidate set of 200 items) that consisted of a context that strongly predicted a specific word, followed by a critical word and a sentence-final word (see Fig. 1). In the *predictable* condition, the critical word was the predictable word. In the *form* condition,

Fig. 1. Results from Experiment 1 (500 ms SOA). ERPs elicited by each condition at Pz in across all items (top panel), in medium-cloze items (left lower panel) and in high-cloze items (right lower panel). Scalp distributions of the ERP effects (implausible condition minus predictable condition) in the N400 time window and LPC time window are shown on the right in each panel.

the critical word was phonologically and orthographically related to the predictable word. The overlap could occur at word-onset (*card-cart*, 15% of the items), word-offset (*luck-duck*, 53.8%), or both (*age-ace*, 23.8%), or involved single-letter addition (*air-hair*, 5.6%), or single-letter deletion (*cold-old*, 1.9%). In the *semantic* condition, the critical word was semantically related to the predictable word. In the *unrelated* condition, the critical word was not related in terms of form or meaning to the predictable word.

We validated our items in four ways. In a cloze probability pre-test, 36 further participants completed each of the context fragments from our candidate set (e.g., *The student is going to the library to borrow a -*) with the first word that came to mind. We excluded items if the predictable word was not the most frequent completion or if it had a cloze probability of less than 30%. Selected items had a mean cloze value of 80% (range 31–100%; see Fig. 1 for example items; the full set of items with cloze values and plausibility ratings are in Supplementary material). We then added an additional word to each item so that ERP responses to the critical words would not be affected by sentence wrap-up.

A further 48 native English speakers judged plausibility of the sentences excluding the post-target word on a scale from 1 (completely implausible) to 5 (completely plausible) for 173 candidate items, together with 64 further sentences that were designed to be plausible. The candidate items were placed in four lists, each containing one version of each item and 43 or 44 sentences from each condition. We excluded items in which the predictable condition had a mean plausibility rating below 3.5 or any other condition had a mean plausibility rating over 3. For the remaining 160 items, the semantic condition was more plausible than the form condition (*Mean Difference* = .16, *SD* = .56), $t(159)$ = 3.66, p < .001, or the unrelated condition (*Mean Difference* = .15, *SD* = .52), $t(159)$ = 3.75, p < .001, whereas the form and unrelated conditions were equally implausible.

Lexical characteristics of the critical words in each condition are shown in Table 1. We evaluated form similarity by computing the Levenshtein distance from the predictable word (the minimum number of single-letter edits including addition, deletion, and substitution needed to transform one word into the other). In the semantic condition, the critical word was semantically related to the predictable word. We assessed this similarity by pairwise Latent Semantic Analysis (LSA) (Landauer & Dumais, 1997). In the unrelated condition, the critical word was related neither in form nor in meaning to the predictable word, relative to the form or semantic conditions. Among the non-predictable conditions, the form condition had a smaller Levenshtein distance to predictable words than the semantic condition, $t(159)$ = −29.4, p < .001, and the unrelated condition, $t(159)$ = −37.1, p < .001. The semantic condition had a larger Levenshtein distance than the unrelated condition, $t(159)$ = 5.2, p < .001. The direction of this difference means that any effect of semantic similarity could not in fact be due to form similarity. The semantic condition had a higher LSA than the form condition, $t(158)$[4] = 26.1, p < .001, and the unrelated condition, $t(159)$ = 27.2, p < .001. The form condition had a higher LSA than the unrelated condition, $t(158)$ = 2.4, p < .05. As will become clear in our Results section, the difference in Levenshtein distance between the semantic and the unrelated condition, and the LSA difference between the form condition and the unrelated condition, cannot explain our results.

Procedure

The 160 sentences were divided into four counterbalanced lists so that each list contained only one condition per sentence, but that across the four lists each condition for each sentence occurred equally often. They were combined with 64 additional plausible filler sentences and presented in the same randomized order for every participant with the constraint that no more than three items from the same condition appeared consecutively. Each participant thus saw a total of 104 plausible and 120 implausible sentences.

Participants silently read sentences from a computer display, presented word by word at a regular pace (300 ms word duration, 200 ms inter-word interval; sentence final words had a 600 ms duration). A fixation-cross followed each sentence, at which point participants could start the next sentence by a button-press. Yes–No comprehension questions appeared on 25% of the trials (mean accuracy across participants, M = 96.1%, SD = 3.6, range 86.0–100%, 6.6% of the responses are excluded due to time-outs). The experiment took approximately 40 min.

Electroencephalogram (EEG) recording and data processing

The electroencephalogram (EEG) was recorded at a sample rate of 512 Hz and with 24-bit AD conversion using the Biosemi ActiveTwo system (BioSemi BV, Amsterdam, The Netherlands). This system's hardware is completely DC coupled and applies digital low pass filtering through its ADC's decimation filter (the hardware bandwidth limit), which has a 5th order sinc response with a −3 dB point at 1/5th of the sample rate (i.e., approximating a low-pass filter at 100 Hz). Data was recorded from 64 EEG, 4 EOG, and 2 mastoid electrodes using the standard 10/20 system (for details, see Nieuwland, 2014). Offline, the EEG was re-referenced to the mastoid average and filtered further (0.019–20 Hz plus 50 Hz Notch filter). Data was segmented into 1200 ms epochs (−200 to 1000 ms relative to critical word onset), corrected for eye-movements using the Gratton and Coles regression procedure as implemented in BrainVision Analyzer (Brain Products ©), baseline-corrected to −100 to 0 ms, automatically screened for movement- or electrode-artefacts (minimal/maximal allowed amplitude = −75/75 μV), and averaged per condition per participant. The mean number of artefact-free trials per condition was 37, with no difference across conditions.

Statistical analysis

Mean amplitude was computed per condition at 16 EEG electrodes (F1/F3/FC1/FC3/CP1/CP3/P1/P3 plus right-hemisphere equivalents), in the N400 time window (350–450 ms) and the LPC time window (600–1000 ms). Effects of condition and scalp distributions effects were tested with a 4 (Condition: Predictable, Form, Semantic, Unrelated) by 2 (Hemisphere: left, right) by 2 (Anteriority: Frontal–Central, Central–Parietal) repeated-measures ANOVA. When appropriate, Greenhouse–Geisser corrections and corrected F-values are reported. Only statistical results with p < .1 are reported. Additionally, we divided the items into high- and medium-cloze probability sets to test an effect of cloze probability with a condition by cloze ANOVA, focusing on relevant conditions.

Results

Visual inspection of the data indicates that all implausible conditions elicited larger N400s than the predictable condition (see Fig. 1). These N400 effects were widely distributed and visible at most channels. Figures showing all channels are in Supplementary material. The form condition also showed a post-N400 enhanced positive deflection

---

[4] One item in the form condition did not yield an LSA value, which is why the comparisons involving this condition have 158 degrees of freedom.

Table 1
Lexical characteristics of critical words (*SD*s in parentheses).

Condition	Length	Frequency	Phonological density	LSA	Levenshtein distance	Concreteness	Familiarity
Predictable	4.4 (1.1)	4.8 (0.5)	8.5 (6.0)	–	–	510 (108)	566 (40)
Form	4.5 (1.1)	4.3 (0.7)	9.0 (6.1)	0.07 (0.1)	1.2 (0.5)	497 (105)	526 (52)
Semantic	5.3 (1.7)	4.3 (0.7)	5.2 (5.5)	0.46 (0.2)	4.8 (1.6)	494 (101)	546 (50)
Unrelated	4.4 (1.0)	4.4 (0.6)	7.0 (6.43)	0.06 (0.1)	4.1 (1.0)	502 (108)	539 (45)

The word frequency was taken from Subtlex (http://zipf.ugent.be/open-lexicons/interfaces/subtlex-uk/). Phonological density represents orthographic neighbourhood size from MCWord (http://www.neuro.mcw.edu/mcword/). LSA shows the results of pair-wise comparison of semantic similarity scores between predictable words and words in each condition (http://lsa.colorado.edu/). Distance represents Levenshtein distance from corresponding predictable words. Concreteness and familiarity ratings are taken from MRC Psycholinguistic Database (http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm).

Table 2
Pairwise *t*-test results for Experiment 1 (500 ms SOA) on mean ERP amplitude per condition at posterior channels in the N400 350–450 ms time window (top-right half) and in the LPC 600–1000 time window (bottom-left half).

Time window		N400			
	Condition	Predictable	Surface	Semantic	Unrelated
LPC	Predictable		-4.8 (3.2) -10.2***	-3.7 (3.1) -8.4***	-5.0 (3.8) -9.2***
	Form	-1.2 (2.9) -3.0**		1.1 (2.2) 3.3**	-0.2 (3.0) -0.5
	Semantic	0.006 (2.9) 0.01	1.2 (3.3) 2.6**		-1.3 (2.4) -3.7***
	Unrelated	0.4 (3.4) 0.9	1.6 (2.6) 4.4***	0.4 (2.3) 1.2	

Note: The values in each cell correspond to the mean voltage difference (row condition values were subtracted from column conditions); *SD* (in parentheses); *t*-value (*df* = 47); significance level, represented as
* <.1.
** <.05.
*** <.001.

compared to the other conditions, starting from about 600 ms and lasting until about 1000 ms, which was most prominent at posterior channels.

The N400 time window

N400 analysis revealed a significant effect of condition, $F(3,69) = 24.6$, $MSE = 14.5$, $p < .001$, and a significant interaction of condition by anteriority, $F(2.1,48.1) = 11.3$, $MSE = 1.9$, $p < .001$, indicating that effects of condition were more robust at posterior channels, $F(3,69) = 29.6$, $MSE = 4.4$, $p < .001$, than anterior channels, $F(3,69) = 16.0$, $MSE = 3.6$, $p < .001$. For further analysis, we therefore performed pairwise comparisons between conditions at the posterior channels where N400 modulations were largest (see Table 2, top-right cells). All three non-predictable conditions elicited larger (more negative) N400s than the predictable condition. Critically, the semantic condition elicited reduced N400s compared to the unrelated conditions.

The late positivity time window

This analysis revealed a significant condition by anteriority interaction, $F(3,69) = 9.9$, $MSE = 1.1$, $p < .001$, a marginally significant interaction of condition by hemisphere, $F(2.2,50.8) = 2.6$, $MSE = 0.8$, $p = .08$, and a marginally significant three-way interaction of condition by anteriority by hemisphere, $F(3,69) = 2.3$, $MSE = 0.3$, $p = .09$.

Since previously reported LPC has been largest at posterior channels, we followed up on the condition by anteriority interaction with one-way ANOVAs at anterior and posterior channels separately. The effect of condition was marginally significant at posterior channels only, $F(2.4,55.6) = 3.0$, $MSE = 5.1$, $p = .05$, and was not significant at anterior channels, $F < 1.6$. Table 2 lists the follow-up pairwise comparisons performed at posterior channels (bottom-left cells). Form-related words elicited an enhanced positivity compared to all other conditions.

Effects of cloze probability

We tested whether the observed N400 and LPC modulations were dependent on the cloze probability of predictable words. To do this, we compared effects in high-cloze and medium-cloze items. We used a median split ($Mdn = 86$) to form a high-cloze subset (83 items, cloze $M = 93.5$, $SD = 4.7$) and a medium-cloze subset (77 items, cloze $M = 65.1$, $SD = 15.3$). Importantly, the two subsets did not differ in plausibility ratings, in frequency, in word length, in Levenshtein distance, in word LSA nor

in context LSA, all *F*s < 1.1 (for details, see Table 4). ERP waveforms for high-cloze and medium-cloze items separately are shown in Fig. 1.

We used only the posterior channels, where, consistent with the findings in previous literature (e.g., Kim & Lai, 2012; Laszlo & Federmeier, 2009), N400 and LPC effects had been maximal. For the semantic prediction reduced-N400 effect, we tested the effect of cloze value on the crucial difference between the semantic condition and the unrelated condition. A 2-way condition by cloze ANOVA revealed a significant effect of condition, $F(1, 23) = 8.4$, $MSE = 6.6$, and of cloze, $F(1, 23) = 9.0$, $MSE = 7.1$, ps < .05, and a marginally significant interaction of condition by cloze, $F(1, 23) = 4.0$, $MSE = 3.5$, $p = .06$. Follow-up *t*-tests comparing the semantic and the unrelated conditions revealed that the N400 reduction was robust in the high-cloze subset ($M = 2.3$ µV, $SD = 3.3$), $t(47) = 4.9$, $p < .001$, but not in the medium-cloze subset ($M = .75$ µV, $SD = 3.2$), $t(47) = 1.6$, $p = .1$. Absence of N400 reduction for the form condition (relative to the unrelated condition) was observed in the high-cloze and medium-cloze subsets alike, ps > .3.

For the form prediction LPC analysis, all four conditions were included to test whether the LPC effect was observed only for the form condition. A condition by cloze ANOVA revealed a significant effect of condition, $F(3, 69) = 4.0$, $MSE = 9.2$, $p < .05$, a marginally significant main effect of cloze, $F(1, 23) = 4.2$, $MSE = 10$, $p = .05$, and a significant interaction of condition by cloze, $F(2.1, 48.9) = 4.5$, $MSE = 9.6$, $p < .05$. Follow-up pairwise comparisons showed that, for high-cloze items, the form condition elicited an LPC effect compared to the predictable condition ($M = 3.4$ µV, $SD = 4.9$), $t(47) = 4.8$, $p < .001$, semantic condition ($M = 2.8$ µV, $SD = 4.7$), $t(47) = 4.1$, $p < .001$, and unrelated condition ($M = 2.6$ µV, $SD = 5.4$), $t(47) = 3.4$, $p = .001$. In contrast, for medium-cloze items, the unrelated condition showed a negative going shift relative to predictable condition, $t(47) = -3.8$, $p < .001$, form condition, $t(47) = -2.8$, $p < .05$, and semantic condition, $t(47) = -2.1$, $p < .05$.

Discussion

We investigated whether readers pre-activate semantic features and (phonological/orthographic) form features in a high-cloze sentence context where strong lexical predictions can be made. Critical words that were semantically related to high-cloze target words elicited a diminished N400 effect compared to unrelated words. Form-related words showed no N400 reduction, but elicited a post-N400 enhanced positivity (posterior LPC) relative to other conditions. The N400 result suggests that participants pre-activated semantic but not form information, whereas the LPC effect suggests that participants detected the form similarity with predictable words. Both effects were robust only in high-cloze sentences, even though the medium-cloze and high-cloze sentences were matched on plausibility and other relevant variables. We consider this strong evidence that both effects arise from prediction of target word meaning. However, one remaining question from Experiment 1 is whether form pre-activation never occurs, or whether it occurs under some experimental conditions but not others.

Experiment 2

Experiment 1 did not generate evidence of form pre-activation. It is possible that readers simply do not pre-activate the form of predictable words. However, it is also possible that form is pre-activated only when time or resources allow, as production-based prediction accounts predict. Do readers pre-activate form when they have more time to generate predictions, for example, when sentences are presented at a slower rate? Experiment 2 tested this possibility by increasing the SOA between words from 500 ms to 700 ms. A 700 ms SOA has also been used as a long-SOA condition in an investigation of presentation rate on prediction during sentence processing (Dambacher et al., 2012) and in word-priming studies (Hill, Ott, & Weisbrod, 2005; Luka & Van Petten, 2014).

There is some reason to believe an increased SOA might enhance pre-activation. Several previous studies have showed an effect of SOA on processing associated with prediction. However, to date, SOA manipulations appear to have been chiefly investigated with respect to semantic prediction. For example, SOA has a clear impact on semantic priming N400 effects. Semantic priming N400 effects suggest that people activate a set of words that are associated with the prime word, which facilitates the processing of the target word (Roland, Yun, Koenig, & Mauner, 2012). Hill et al. (2005) reported a larger N400 priming effect for a longer SOA (700 ms) than a shorter SOA (150 ms), and suggested that the longer SOA led to deeper semantic processing. Luka and Van Petten (2014) presented participants pairs of words that were strongly, moderately, or weakly associated with each other, either simultaneously or with a 700 ms SOA. They found that stronger semantic association was associated with smaller N400s. Critically, this N400 effect was delayed in the simultaneous presentation, especially for strongly associated word pairs, suggesting that more time enhances pre-activation via semantic association (i.e., priming). These studies investigated the effect of SOA only on word-to-word semantic relatedness effects, and it is unclear whether similar effects occur during sentence reading when predictions for a specific word can be made based on more constraining contextual information.

Kutas (1993) had participants read sentences that ended with the highest cloze word, an incongruent word that was semantically related to the highest cloze word, or an incongruent word that was semantically unrelated to the highest cloze word, at four different SOAs (100, 250, 700, or 1150 ms). N400s were smaller for semantically related incongruent words than for semantically unrelated incongruent words at all SOAs. More interestingly, the N400 peak latency difference was delayed for the fastest 100 ms SOA compared to the slower (250–1150 ms) SOAs, and the N400 difference effect was smaller for the 100–250 SOAs than for the 700 ms SOA (albeit not smaller than for the 1150 ms SOA). The results suggest that the effect of semantic relatedness to the expected words is very robust and occurs at relatively slow or fast presentation rates.

Dambacher et al. (2012) found N400 effects for low-predictable words relative to high-predictable words in three different SOAs (280 ms, 490 ms, and 700 ms). The N400 amplitude difference was the smallest in the 280 ms SOA, with no difference between 490 ms and 700 ms SOAs, and the onset latency of the N400 effect was delayed in the 280 ms SOA relative to the 700 ms SOA. However, their high-predictable words were also more plausible the low-predictable words, so it remains unclear whether the effects of SOA on the N400 reflect the semantic processes associated with pre-activation or plausibility.

In replication of Federmeier and Kutas (1999), Wlotko and Federmeier (2015) found that N400 effects were reduced for words semantically related to a predictable word relative to unrelated words. This N400 reduction was found at a 500 ms SOA, but not at a 250 ms SOA. Along with Dambacher et al. (2012), the impact of faster presentation on the N400 effects points to the possibility that prediction-related N400 effects may suffer from uncomfortably rapid serial presentation rates.

These studies suggest that manipulations of timing may have some effects on pre-activation,[5] but there have been no investigations of the effects of timing on form pre-activation. In our Experiment 1, participants pre-activated meaning but not form, as indicated by a reduced N400 for semantically related words relative to unrelated words, but not a reduced N400 for form-related words relative to unrelated words. Experiment 2 used a longer SOA and examined whether form pre-activation occurs when there is more time available during sentence comprehension to generate predictions. If form features are pre-activated when people read at this slower presentation rate, we would expect to see reduced N400s for form-related words relative to unrelated words.

Methods

Participants

Twenty-four English monolinguals (6 males and 18 females, age $M = 20.3$ years, $SD = 2.9$) took part in the experiment, having given informed consent. The participants were from the same population as in Experiment 1, but had not participated in Experiment 1. All participants were right-handed and free from neurological or language disorders.

Stimuli and experimental design

The stimuli and experimental design were identical to those used in Experiment 1.

Procedure

The procedure was identical to Experiment 1, except that the SOA was changed to 700 ms (500 ms word duration, 200 ms inter-word interval; sentence-final words

had an 800 ms duration). Mean accuracy for comprehension questions was 90.8% ($SD = 3.9$, range 83.9–96.4%).

Electroencephalogram (EEG) recording and data processing

The data were processed in the same way as in Experiment 1. The mean number of artefact-free trials in Experiment 2 per condition was 36, with no difference across conditions.

Statistical analysis

The same statistical analysis was conducted as for Experiment 1.

Results

Visual inspection of the data indicates that all implausible conditions elicited larger N400s than the predictable condition, and the N400 effect was reduced for the semantic condition (see Fig. 2), as had also been observed in Experiment 1. The form condition elicited the largest LPC at posterior channels among all the conditions.

The N400 time window

The analysis revealed a significant effect of condition, $F(3, 69) = 18.5$, $MSE = 17.2$, $p < .001$, and a significant interaction of condition by anteriority, $F(3, 69) = 4.3$, $MSE = 1.8$, $p < .05$, which was due to a stronger effect of condition at posterior channels, $F(3, 69) = 19.2$, $MSE = 5.0$, $p < .001$, than at anterior channels, $F(3, 69) = 15.0$, $MSE = 4.5$, $p < .001$. Following the same analysis steps in Experiment 1, pairwise comparisons were performed at posterior channels. All the implausible conditions elicited larger N400s than the predictable condition, but the semantic condition elicited reduced N400s relative to the form and the unrelated conditions (see Table 3, top-right cells).

The late positivity time window

The analysis revealed a significant interaction of condition by anteriority, $F(3, 69) = 24.1$, $MSE = 1.1$, $p < .001$, which was driven by the fact that the effect of condition was significant at posterior channels, $F(2.4, 55) = 4.7$, $MSE = 5.1$, $p < .05$, but not at anterior channels, $F < 1$. Similarly to Experiment 1, the pairwise comparisons at posterior channels revealed that the form condition elicited enhanced positivity relative to all the other conditions (see Table 2, bottom-left cells).

Effects of cloze probability

We compared high-cloze and medium-cloze items using the same median split as in Experiment 1 in order to test the effect of predictability. For the N400 window, we performed a 2-way condition by cloze ANOVA at posterior channels including the form, semantic, and unrelated conditions in order to allow investigation of pre-activation of both form and meaning. The analysis revealed a significant interaction of condition by cloze, $F(1.6, 37) = 6.0$, $MSE = 9.0$, $p < .05$, which arose from the effect of condition being significant in the high-cloze items, $F(2, 46) = 5.8$, $MSE = 9.3$, $p < .05$, but not in the medium-cloze items, $F < 1.5$. For the high-cloze items,

---

[5] An important caveat to an interpretation of SOA-based peak latency differences, however, is that at short SOAs the N1-P2 ERP complex elicited by the subsequent word occurs in the N400 time range of the critical word. Through component overlap, ERPs elicited by the subsequent word at short SOAs (but not at a 500 ms or slower SOA) can thus 'cut short' the N400 component before it reaches its full peak.
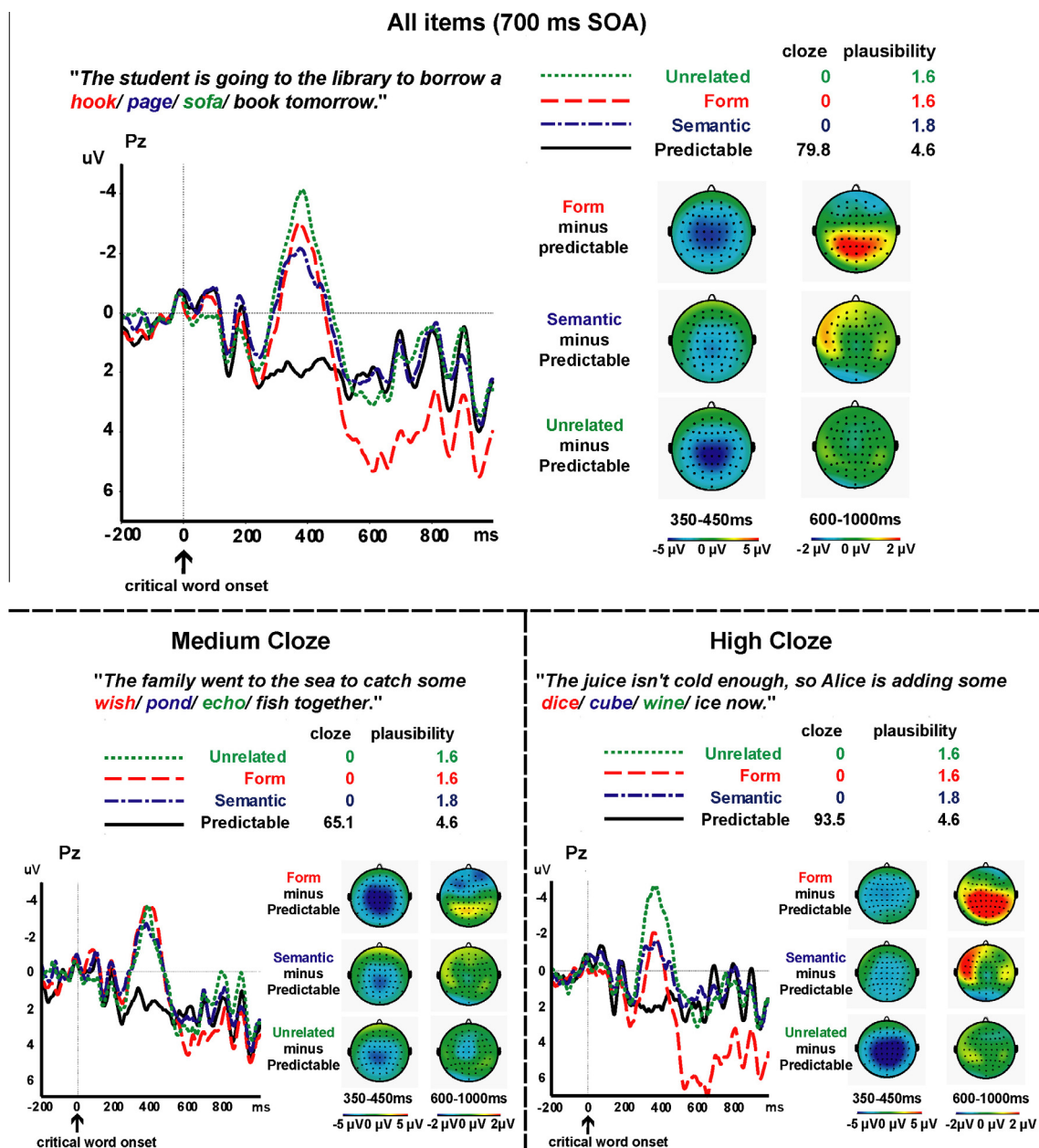
Fig. 2. Results from Experiment 2 (700 ms SOA). ERPs elicited by each condition at Pz in across all items (top panel), in medium-cloze items (left lower panel) and in high-cloze items (right lower panel). Scalp distributions of the ERP effects (implausible condition minus predictable condition) in the N400 time window and LPC time window are shown on the right in each panel.

N400s were reduced both for the form condition, ($M = 2.6$ μV, $SD = 3.6$), $t(47) = 5.0$, $p < .001$, and for the semantic condition, ($M = 2.7$ μV, $SD = 4.6$), $t(47) = 4.0$, $p < .001$, relative to the unrelated condition (see Fig. 2, right lower panel).

For the LPC analysis, a 2-way condition by cloze ANOVA at posterior channels with all the conditions revealed a significant effect of condition, $F(2.4, 54) = 5.4$, $MSE = 10.3$, $p < .05$. However, neither the effect of cloze nor the interaction of condition by cloze was significant, $Fs < 1.6$.

Between-experiment comparisons: effects of SOA

The critical difference between the results of the two experiments was the presence of the N400 modulation for form-related words in high-cloze items in Experiment 2 but not in Experiment 1. To specifically test this effect, we conducted a between-experiment comparison at posterior channels in high-cloze items, using a difference value between the form and the unrelated conditions. An independent samples t-test revealed a significant effect of SOA, $t(94) = -2.5$, $p < .05$. In the LPC time window, an

Table 3
Pairwise *t*-test results for Experiment 2 (700 ms SOA) on mean ERP amplitude per condition at posterior channels in the N400 350–450 ms time window (top-right half) and in the LPC 600–1000 time window (bottom-left half).

Time window		N400			
	Condition	Predictable	Surface	Semantic	Unrelated
LPC	Predictable		-4.0 (3.6) -7.6***	-3.1 (2.6) -8.4***	-4.4 (3.5) -8.9***
	Form	-1.8 (3.5) -3.6***		0.81 (2.8) 2.0**	-0.47 (3.1) -1.1
	Semantic	-0.02 (2.9) 0.06	1.8 (3.2) 3.9***		-1.3 (3.5) -2.6**
	Unrelated	-0.04 (2.5) -0.1	1.76 (2.1) 5.7***	-0.06 (2.9) -0.14	

Note: The values in each cell correspond to the mean voltage difference (row condition values were subtracted from column conditions); *SD* (in parentheses); *t*-value (*df* = 47); significance level, represented as
* <.1.
** <.05.
*** <.001.

independent samples *t*-test comparing the same 2 conditions showed no significant effect of SOA, *p* = .8.

Discussion

Experiment 2 investigated whether pre-activation of form features would occur when sentences are presented at a slower presentation rate (700 ms SOA) than in Experiment 1 (500 ms SOA). Experiment 2 partially replicated Experiment 1; semantically related words elicited reduced N400s compared to unrelated words, and form-related words elicited an enhanced later positivity relative to all the other conditions. However, unlike in Experiment 1, form-related words showed an N400 reduction for high-cloze items, and elicited an LPC effect in medium-cloze items as well as in high-cloze items. The results suggest that participants pre-activated the forms of highly predictable words, but not of moderately predictable words. The results also suggest that our participants detected conflict between form-related words and predictable words, irrespective of whether the form features were pre-activated or not.

General discussion

Two ERP experiments examined pre-activation of form and meaning during sentence reading, and whether pre-activation of semantic and form features depends on the time constraints on reading. Participants read high-cloze sentences that were completed with the predictable word, an anomalous word that was either semantically related or form-related to the predictable word, or an unrelated word. The rate of the word-by-word presentation was 500 ms in Experiment 1 and 700 ms in Experiment 2. Anomalous words in all conditions elicited an N400 effect compared to the predictable word, but, at both SOAs,

N400s for semantically related words were reduced compared to unrelated words. In contrast, form-related words elicited reduced N400s only at the 700 ms SOA, and only in the high-cloze item subset. However, form-related words elicited an enhanced post-N400, Late Positive Component (LPC) at both SOAs. This LPC effect occurred irrespective of whether form-related words elicited reduced N400 effects, and was elicited only by high-cloze items at the 500 ms SOA, but by medium-cloze items and high-cloze items alike at the 700 ms SOA. The main novel contributions of this work are that (1) both meaning and form can be pre-activated, but pre-activation of form is more influenced by time constraints than pre-activation of meaning, and (2) whether or not the form of predicted words is pre-activated, form similarity to predicted words incurs additional post-N400 processing costs, suggestive of an interpretation conflict between expected input and encountered input.

Pre-activation of semantic features

The results of both our experiments strongly suggest that lexical prediction entails the pre-activation of semantic features, by ruling out the effects that could be associated with an account in terms of ease of integration (see Federmeier & Kutas, 1999, for discussion). Under an integration account, the facilitation of semantically related words, reflected in the reduced N400, occurs primarily because these words are more plausible sentence continuations. Although semantically related words in the complete set of items were rated as slightly less implausible than the form/unrelated conditions (1.8 compared to 1.6, on a plausibility scale of 1 to 5), the currently observed N400 reduction for semantically related words cannot straightforwardly be explained in terms of plausibility or other factors (e.g., semantic priming from context, lexical

Table 4
Lexical characteristics of critical words in high- and medium cloze item sets. For these variables, the only robust difference between high and medium cloze sets (pair-wise *t*-tests) was found for unrelated words, which had higher LSA values in the medium cloze set than in the high cloze set, $t(142.6) = -2.3$, $p < .05$.

Condition	Cloze set	Length	Frequency	Phonological density	LSA	Distance	Context LSA	Cloze	Plausibility
Predictable	High	4.37	4.83	8.57			0.22	93.50	4.58
	Medium	4.48	4.79	8.40			0.20	65.12	4.59
Form	High	4.51	4.28	8.65	0.07	1.30	0.10	0	1.58
	Medium	4.48	4.28	9.38	0.08	1.18	0.09	0	1.65
Semantic	High	5.23	4.29	5.08	0.48	4.70	0.17	0	1.76
	Medium	5.43	4.40	5.34	0.44	4.88	0.16	0	1.78
Unrelated	High	4.36	4.34	6.90	0.05	4.07	0.09	0	1.61
	Medium	4.51	4.40	7.09	0.07	4.21	0.10	0	1.63

characteristics including frequency and word length). This conclusion is based on the fact that N400 reduction was found only for high-cloze items and not for medium-cloze items, while the high-cloze and medium-cloze items were matched on plausibility and other relevant variables (see Table 4). In other words, the N400 reduction for the semantically related condition was not dependent on plausibility, but depended on the cloze probability of the predictable word. We take this as strong evidence for pre-activation of semantic information, at least in highly constraining sentences.

Our study did not find clear evidence for semantic pre-activation in the medium-cloze sentences (*mean cloze* = 65%). This result is inconsistent with other studies that found an N400 reduction for semantically related words despite relatively low cloze probabilities (Federmeier & Kutas, 1999; Thornhill & Van Petten, 2012). Thus, the semantic relatedness of the related and predictable words may have been stronger in these previous studies than in our study. Our findings suggest that prediction may not always occur, even when cloze probability is relatively high. This may be incompatible with models which regard prediction as a fundamental aspect of language processing (e.g., Dell & Chang, 2014). However, if engagement of online prediction varies during language processing, it may be likely that factors besides cloze predictability (i.e., semantic relatedness, experimental design factors such as the inclusion of nonwords and differences in task instructions) may affect the detection of online prediction. Furthermore, it could be the case that the pre-activation of particular features during prediction is highly contextually mediated, such that the system only pre-activates those features that are relevant or strongly associated with the particular situation that unfolding evidence supports (Metusalem et al., 2012).

Semantically related words elicited identical N400 reductions at shorter and longer SOAs, suggesting that longer SOA did not lead to stronger pre-activation of semantic features. This does not necessarily mean that semantic pre-activation is an automatic process that is

unaffected by time constraints. Our results are consistent with those reported by Dambacher et al. (2012), where SOAs of 490 ms and 700 ms elicited similar N400 effects of cloze probability. However, they additionally found that when sentences were presented at 280 ms SOA, the N400 effect was smaller than at the other two SOAs, and the onset latency of the N400 effect was delayed compared to the 700 ms SOA. In an earlier study, Kutas (1993) also found a delay in the onset and peak latencies of an N400 anomaly effect when using 100 ms SOA, which was faster than a normal reading speed. Such findings suggest that N400 effects for unexpected or semantically anomalous words can be affected by time constraints (i.e., word presentation rate), and appear to show a delay at relatively short SOAs. The 500 and 700 ms SOAs in the current study may have been too long to generate such patterns.

Pre-activation of form features

While our results show that form features can be pre-activated, form pre-activation depended on the time that was available to generate predictions during reading. At a 500 ms SOA, there was no sign of pre-activation of form features, even in the most constraining sentences (cloze value 94%) at a relatively high level of form-similarity (as reflected by a relatively low Levenshtein distance of 1.3). This result appears inconsistent with previous studies by Laszlo and Federmeier (2009) and Kim and Lai (2012), who found facilitation effects for pseudowords that were orthographically similar to predictable words using 500 ms SOA and 550 ms SOA, respectively. This inconsistency could indicate that readers process real words that resemble predictable words differently from pseudowords or non-words that resemble predictable words.

However, the discrepancy between our findings and those of Laszlo and Federmeier (2009) may also have to do with the different task instructions. Participants in the present study answered comprehension questions after some trials, whereas participants in Laszlo and Federmeier (2009) were asked to judge each sentence on

whether it was a "normal English sentence". This explicit judgment task may have drawn extra attention to the included nonwords and increased the task-relevance of form-related non-words. This would have been exacerbated by the fact that the form-related targets were more similar to the correct words (i.e., had lower Levenshtein distances), than those in our study. Such issues complicate a direct comparison of our findings with those of previous studies.

In our study, form-related words showed facilitation effects at the 700 ms SOA, reflected in an N400 reduction, but this effect was limited to high-cloze sentences. It thus appears that very high predictability is critical for the pre-activation of form features. This explanation fits with related studies: All the reviewed studies that found an N400 modulation for words sharing orthographic or phonological features with predictable words used critical sentences with cloze probability of about 90% (Kim & Lai, 2012; Laszlo & Federmeier, 2009; Vissers et al., 2006). Importantly, Experiment 2 replicated the previously reported N400 reductions using real words only, and without a secondary task that required explicit judgments about the critical words. Moreover, the high-cloze items and medium-cloze items did not differ in form similarity (see Table 4, Levenshtein distance), plausibility, semantic relatedness to preceding contexts or other lexical characteristics that might explain the N400 effect difference. Therefore, we interpret the predictability-dependent facilitation effect for the form-related words as reflecting pre-activation of form features as a consequence of prediction of a specific word.

A comparison of the Experiments 1 and 2 suggests that pre-activation of form features is more dependent on time constraints than pre-activation of semantic features. This finding suggests that pre-activation of semantic features is more robust than pre-activation of form features. Strictly speaking, we cannot rule out that the form-prediction SOA effects occur because form-predictions take slightly longer to develop from the presentation of the pre-critical word. In the latter case, we expect pre-activation of form not to be a function of general SOA but only of the time between the critical word and the pre-critical word. However, rather than making such a strong claim about the absolute time course of the unfolding form-prediction, we think that slower SOAs might in principle benefit all aspects of prediction. After all, people are more likely to finish the sentences of someone who speaks slowly or hesitantly than of someone who speaks fast and fluently (see Gambi & Pickering, 2011).

Pre-activation pattern and production-based prediction accounts

Our findings are compatible with an account in which comprehenders use the production system to make predictions during comprehension (Pickering & Garrod, 2007, 2013; see also Federmeier, 2007). According to this proposal, comprehenders covertly imitate what they are hearing, so that they generate a production-based representation. They then engage some of the mechanisms of language production to predict upcoming words (indicating, roughly, what they would themselves say next at that point). Normally, comprehenders do not have time or resources to construct a full "implemented" representation of what they would say next (and instead construct a forward model, as discussed in Pickering & Garrod, 2013). But when time and resources allow, they run through the stages involved in language production, which involve semantic representations followed by form representations (e.g., Levelt et al., 1999). We therefore propose that comprehenders in Experiment 1 had the time to construct meaning (thus leading to a reduced N400 in the semantic condition) but not form (thus leading to no N400 reduction in the form condition). In contrast, comprehenders in Experiment 2 had the time to construct both meaning and form (thus leading to N400 reductions in both conditions).[6]

If similar effects of meaning and form pre-activation had been obtained at both SOAs, it would have suggested that participants pre-activated a specific lexical item (i.e., lemma) first, wherefrom the activation spread across semantically and form-related lemmas. If this were the case, the pre-activation pattern would have been incompatible with a prediction-with-implementation account.

A caveat to this claim, though, is that we did not have a condition where the critical word was related to the predictable word both in form and meaning. (Because of the very limited number of such lexical pairs, it is unclear whether enough items for ERP signal-to-noise requirements could be constructed.) Even with such a condition, it would be impossible to compare the strength of semantic relatedness and form relatedness, as they are not quantified in the same way. Hence, we cannot ensure that our semantically related words and form related words were equally strongly 'related' to the predictable words.

Furthermore, our observed pattern of results is also consistent with a comprehension system in which activation cascades from the semantic to the form level, regardless of engagement of the production system during prediction. It is possible that participants first pre-activated semantic information, and this activation cascaded to form information, purely within the comprehension system. As this suggests that semantic pre-activation occurs prior to form pre-activation, the effect of SOA in our study can be explained by the assumption that the SOA was slow enough for the cascading to the form level to occur in Experiment 2 but not in Experiment 1. Evidence for such cascaded lexical activation has been found in comprehension (Apfelbaum, Blumstein, & McMurray, 2011; Huettig & McQueen, 2007) as well as in production (Morsella & Miozzo, 2002). Although well-established models of language comprehension and language prediction do not yet clearly formalize the notion of cascaded processing, one might reasonably assume that a cascading architecture in comprehension could work in the following basic way: Activation cascades from word

---

[6] An interesting point is that the form-related LPC did occur in Experiment 1. It may be that some form-related pre-activation did occur in Experiment 1, but it was not ready at the point at which the N400 was elicited, or there was a need for concurrent activation from the form-related target word.

form level to semantic level. In contrast, production models that entail cascading (cf. Levelt, 1999) posit that cascading occurs from semantic level to word form level. Minimally, our results clearly support a cascaded processing architecture, whatever the nature of the representations that are being activated. Given that our data follow the cascaded pattern assumed in production models (i.e., semantic followed by form), we find the directionality of activation as most consistent with the predictions of a production-based prediction account. But our findings do not offer conclusive evidence to the point of exclusion of alternative accounts, and are consistent with both production-based prediction and cascaded lexical activation accounts.

Monitoring and reanalysis processing for form-related words

In both experiments, the form-related condition elicited a post-N400, posterior positive deflection (LPC) that depended on the cloze probability of predictable words. Critically, this LPC effect occurred only for the form condition, in comparison to the predictable words but also to the other two implausible conditions. Therefore, this effect cannot be explained in terms of implausibility under high-constraint conditions, as proposed by Van Petten and Luka (2012) in a review of post-N400 LPCs (see also DeLong, Quante, & Kutas, 2014). We can identify three possible accounts for this effect. According to a monitoring account, form similarity increases monitoring processes and triggers a general reanalysis to check for processing errors (Van de Meerendonk, Kolk, Chwilla, & Vissers, 2009; Van Herten, Chwilla, & Kolk, 2006). According to a misspelling account, participants may have considered the form-related words as being misspellings of the predicted word, triggering a repair of the surface feature that differed between the predicted and encountered input (e.g., Kim & Osterhout, 2005). Importantly, both these accounts assume that comprehenders detect a difference between the encountered input and the input that they predicted, but the accounts differ in terms of whether comprehenders lay the blame on errors in their own comprehension processes or on an error in the written input. A third interpretation is that the LPC effect only reflects the detection of similarity to the predicted form, without comprehenders considering any input or process to be erroneous.

The combination of an N400 effect and subsequent LPC effect suggests that participants in our experiment did not take the form-related words purely as a misspelling. It suggests that semantic information associated with form-related words was indeed accessed. Moreover, we observed the LPC effects whether or not the N400 for form-related words was reduced (i.e., in Experiments 1 and 2), indicating that pre-activation of form was not necessary for the occurrence of the subsequent processes reflected in the LPC. The pre-activation of form information and the detection of form similarity thus appeared to be fairly independent of each other. Detection of form similarity of encountered input with predicted input may thus arise via a bottom-up process of feature activation (Federmeier, 2007), rather than pre-activation.

Conclusions

Current neurobiological accounts of language comprehension assume lexical prediction through pre-activation (Federmeier, 2007; Kutas et al., 2009). However, less is known about how linguistic information is pre-activated. We examined the patterns of co-occurrence of form and meaning pre-activation to test whether the patterns would be consistent with production-based prediction accounts. Our study investigated prediction of form and meaning while participants read grammatical sentences without having to explicitly evaluate the critical words. Anomalous words that were semantically related to predictable words elicited reduced N400 effects compared to unrelated words, and this effect was not influenced by SOA. In contrast, highly predictable form-related words elicited a reduced N400 effect in the slower SOA, suggesting that people pre-activate the semantics of predictable words more strongly than the form. Form-related words also elicited an enhanced, post-N400 posterior positivity at both SOAs, indicating that form similarity between expected and encountered input was detected via a bottom-up mechanism, regardless of whether form features are pre-activated or not. Our results demonstrate that pre-activation of the form of upcoming words depends on the time that readers have to predict, which we suggest is in line with production-based accounts of linguistic prediction.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jml.2015.10.007.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247–264.

Apfelbaum, K. S., Blumstein, S. E., & McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: Evidence for continuous cascading systems. *Psychonomic Bulletin & Review, 18*, 141–149.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*, 67–82.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177–208.

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A. M., & Kliegl, R. (2012). Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia, 50*, 1852–1870.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London, Series B: Biological sciences, 369*, 20120394.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314.

DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language, 121*, 226–239.

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia, 61*, 150–162.

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.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505.

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.

Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. *Frontiers in Psychology, 2*, 1–14.

Hare, M., Elman, J. L., Tabaczynski, T., & McRae, K. (2009). The wind chilled the spectators, but the wine just chilled: Sense, structure, and sentence comprehension. *Cognitive Science, 33*, 610–628.

Hill, H., Ott, F., & Weisbrod, M. (2005). SOA-dependent N400 and P300 semantic priming effects using pseudoword primes and a delayed lexical decision. *International Journal of Psychophysiology, 56*, 209–221.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*, 66–85.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*, 460–482.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.

Kim, A. E., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience, 24*, 1104–1112.

Kim, A. E., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225.

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., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In *Predictions in the Brain: Using Our Past to Generate a Future* (pp. 190–207).

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.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161–163.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

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

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences, 3*, 223–232.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences, 22*, 1–38. discussion 38–75.

Luka, B. J., & Van Petten, C. (2014). Prospective and retrospective semantic processing: Prediction, time, and relationship strength in event-related potentials. *Brain and Language, 135*, 115–129.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

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.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 555–563.

Newman, R. L., & Connolly, J. F. (2004). Determining the role of phonology in silent reading using event-related brain potentials. *Cognitive Brain Research, 21*, 94–105.

Newman, R. L., Connolly, J. F., Service, E., & Mcivor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potentials. *Psychophysiology, 40*, 640–647.

Nieuwland, M. S. (2014). "Who's he?" Event-related brain potentials and unbound pronouns. *Journal of Memory and Language, 76*, 1–28.

Nieuwland, M. S. (2015). The truth before and after: Brain potentials reveal automatic activation of event knowledge during sentence comprehension. *Journal of Cognitive Neuroscience, 27*, 2215–2228.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189–234.

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. A. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience, 8*, 89.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*, 105–110.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences, 36*, 329–347.

Roland, D., Yun, H., Koenig, J. P., & Mauner, G. (2012). Semantic similarity, predictability, and models of sentence processing. *Cognition, 122*, 267–279.

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.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*, 569–613.

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.

Van de Meerendonk, N., Kolk, H. H. J., Chwilla, D. J., & Vissers, C. T. W. M. (2009). Monitoring in language perception. *Linguistics and Language Compass, 3*, 1211–1224.

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.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190.

Vissers, C. T. W. M., Chwilla, D. J., & Kolk, H. H. J. (2006). Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Research, 1106*, 150–163.

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