NeuroImage 189 (2019) 192–201

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Ahead of time: Early sentence slow cortical modulations associated to semantic prediction

Patricia León-Cabrera [a,b,*], Amanda Flores [c], Antoni Rodríguez-Fornells [a,b,d], Joaquín Morís [c,e,**]

[a] Department of Cognition, Development and Educational Psychology, Universitat de Barcelona, Barcelona, Spain
[b] Cognition and Brain Plasticity Unit [Bellvitge Biomedical Research Institute, IDIBELL], Barcelona, Spain
[c] Instituto de Investigación Biomédica de Málaga (IBIMA), Universidad de Málaga, Spain
[d] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
[e] Department of Social Psychology, Universidad de Málaga, Spain

ARTICLE INFO

Keywords:
Semantic prediction
Sentence processing
Slow negative potentials

ABSTRACT

According to prediction-based accounts of language comprehension, incoming contextual information is constantly used to guide the pre-activation of the most probable continuations to the unfolding sentences. However, there is still scarce evidence of the build-up of these predictions during sentence comprehension. Using event-related brain potentials, we investigated sustained processes associated to semantic prediction during on-line sentence comprehension. To address this, participants read sentences with varying levels of contextual constraint one word at a time. A 1000 ms interval preceded the final word, which could be congruent or incongruent. A slow sustained negativity developed gradually over the course of sentences, showing differences across conditions, with increasingly larger amplitudes for high than low levels of constraint. The effect was maximal in the interval preceding the closing word. This interval elicited a left-dominant slow negative potential with a graded amplitude modulation to contextual constraint, replicating previous results in speech comprehension. We argue that these slow potentials index the engagement of cognitive operations associated to semantic prediction. In addition, we replicated the finding of an earlier onset of the N400 effect (incongruent minus congruent) for high relative to low contextual constraint, suggesting facilitated processing for contextually-supported and highly expected words. Altogether, these results are consistent with prediction-based models of language comprehension and they also strengthen the value of investigating slow components as potential indices of mechanisms linked to language prediction.

1. Introduction

Sentences unfold linearly in time. As every word is encountered, it is incorporated into a broader representation to achieve comprehension rapidly (Van Berkum et al., 1999; Van Petten et al., 1999), in line with our everyday experience of understanding language on the fly. This immediacy might be bolstered by predictive mechanisms, whereby top-down information guides the pre-activation of features associated to the most probable upcoming words to facilitate their processing upon receipt (for a review, Kutas et al., 2011). In the past years, extensive research has shown that words that are more predictable given a previous context show facilitated processing over less predictable ones, as indexed by reduced amplitudes in the N400 component (e.g. Van Petten et al., 1999; Federmeier, 2007; Kutas et al., 2011), and there is evidence that, under some circumstances, readers may pre-activate specific lexical units (Wicha et al., 2003; DeLong et al., 2005; Van Berkum et al., 2005; Freunberger and Roehm, 2016; but see Nieuwland et al., 2018). Conversely, other authors have argued that, given the infinite number of potential continuations to the same phrase, prediction is an unviable strategy in language and that comprehension proceeds in a strictly bottom-up, stimulus-driven fashion (e.g., Jackendoff, 2003, 2007; Morris, 2006). More recently, some researchers have called into question whether prediction is actually necessary for language comprehension (see Huettig and Mani, 2016). To contribute in the quest for

* Corresponding author. Department of Cognition, Development and Educational Psychology, Faculty of Psychology, University of Barcelona, Barcelona, 08035, Spain.
** Corresponding author. Department of Social Psychology, Social Anthropology, Social Work and Social Services, School of Psychology, University of Málaga, Málaga, Málaga, 29017, Spain.
E-mail addresses: leoncabrera.patricia@gmail.com (P. León-Cabrera), joaquin.moris@gmail.com (J. Morís).

https://doi.org/10.1016/j.neuroimage.2019.01.005
Received 9 October 2018; Received in revised form 28 November 2018; Accepted 4 January 2019
Available online 6 January 2019
1053-8119/

understanding better the role of predictive processing in language, the present study investigated electrophysiological correlates associated to the build-up of predictions over the course of sentence comprehension.

In order to investigate the prediction process *per se*, one strategy is to focus on the anticipatory stage of the words that are being predicted. To do so, sentences are built so that their semantic content establishes a weak or a strong expectation for the final word. Highly constraining sentences (HC) have a single best completion (e.g. "the dentist proceeded to clean her … teeth"), whereas low constraining (LC) sentences have more than one likely continuation (e.g. "the meeting was arranged for the … morning/afternoon/evening). In a recent ERP study using this type of paradigm with spoken sentences (León-Cabrera et al., 2017), we inserted a 1-s delay between the penultimate and the final word, thus substantially extending the anticipatory period. We reported the development of a slow negative potential in the anticipatory period prior to closing words, with larger amplitudes for HC than for LC and non-semantic sentences. Interestingly, this neural signature was reminiscent of the stimulus-preceding negativity (SPN), a slow potential that has been described in other domains as an index of anticipation for upcoming relevant events (for reviews, see Brunia et al., 2011; Hackley et al., 2014). More specifically, it showed a similar gradual amplitude increase over time, a graded modulation that agrees with the predictability of the upcoming stimulus and a frontal topographical distribution. Given its characteristics and the task that was employed, we suggested that the slow potential might be a language-related SPN reflecting the semantic anticipation of the upcoming word. In a similar token, Grisoni et al. (2017) reported a slow negative potential prior to the final words for HC (but not LC) sentences. They interpreted it as an index of specific semantic predictions, given that it emerged over the motor areas specifically associated to the upcoming concept –dorsolateral for hand-related verbs (e.g. "write") and ventral for face-related verbs (e.g. "talk").

The investigation of brain activity in the anticipatory period of words has also proven fruitful using different methodological approaches (Dikker and Pylkkänen, 2013; Piai et al., 2014; Rommers et al., 2017; Wang et al., 2018), generally supporting the idea that the information from the context changes the state of the cognitive system before critical words are encountered. However, exploring the anticipatory period of a single word offers a very restricted time interval that can only reveal rather transient processes. According to interactive models of language comprehension, contextual information has an immediate and continuous impact on bottom-up processing (e.g. Nieuwland and Van Berkum, 2006). Therefore, the effect of semantic constraint should be detectable much earlier in time and be captured by sustained, rather than punctate brain processes. Hereafter, we will use the labels "sentence-level" and "local" to distinguish between these periods.

Interestingly, ERPs are well-suited to track the time-course of sentence-level dynamics. In the past, several ERPs studies have found cumulative, sentence-level slow modulations that are distinguishable from faster, word-related processes (for a review, Kutas, 1997). Most of the available multiword ERP studies focused on sentences of varying syntactic complexity. In these, slow sustained positivities were found for sentences with simpler structures which are easy to integrate (Kutas and King, 1996), whereas negativities were associated to increased working memory (WM) demands when processing more complex sentences (in reading comprehension, King and Kutas, 1995; Fiebach et al., 2002; in speech comprehension, Müller et al., 1997). The involvement of WM was further backed-up by a correlation between the amplitude of the slow sustained negativity and individual differences in WM capacity – participants with a lower WM span showed larger negativities than those with a higher span (see also, Vos et al., 2001). On the other hand, WM differences not associated to syntactic structure can also result in sustained negativities. As such, Münte et al. (1998) found that the amount of conceptual information activated by a single word led to a similar modulation.

With regards to the current goal, a previous study compared sentence-level dynamics between different levels of semantic constraint, and did not find differences (Kutas et al., 1988). Critically, however, the experiment was originally designed to investigate the N400 component (see, Kutas and Hillyard, 1980) and thus there are several methodological issues (e.g. averaging sentences with different word lengths) that might have precluded the observation of significant slow wave effects. Using an appropriate task design, we hypothesized that the effect of semantic constraint would be captured by slow sustained modulations, arising through sentence comprehension. Furthermore, we expected to find the same pattern of amplitude differences as that of local slow negative potentials associated to semantic anticipation during spoken sentences (León-Cabrera et al., 2017). Such a correlate would be consistent with prediction-based models of sentence comprehension, whereby the probabilities of upcoming continuations are computed and updated gradually as the semantic context accrues (for a review, Kutas et al., 2011).

The goal of this study was to investigate sustained ERP correlates associated to the build-up of predictions over the course of sentence comprehension. To address this, we adapted the task of a previous study using spoken sentences (León-Cabrera et al., 2017) for visual presentation, which also allowed us to answer additional questions. First, we presented sentences using rapid visual serial presentation (RVSP) to equate the duration of words and track word-by-word changes by temporally aligning them. Critically, this feature of the design allowed a simple and straightforward analysis of the EEG activity during sentence processing. This analysis was not possible in the previous study, given that the stimuli were presented auditorily and, consequently, the sentences had variable durations. Second, the use of a different input modality (i.e. visual) allowed us to test the generality of the findings in the previous study. That is, whether similar results regarding the local negativity and the N400 component would replicate in the visual modality.

All the other features of the task and the materials employed were the same as in León-Cabrera et al. (2017). The sentences had different levels of semantic constraint (HC, LC or none) and the final word appeared after a delay of 1000 ms, in order to extend its anticipatory period. We inserted the delay in the attempt to (1) replicate previous findings of local slow potentials time-locked to the onset of the anticipatory period, and, in case we found a sentence-level modulation (2) determine whether both levels of analysis (i.e. sentence-level and local), captured the same cognitive operations. Based on our previous study, we expected more negative amplitudes for HC than LC contexts. Finally, to assess the potential consequences of prediction and misprediction, the final word turned out to be incongruent in half of the trials so as to evoke the N400 component. If some representation of the final words become pre-activated at some point, an earlier onset of the N400 effect in the HC condition could be observed, as an index of processing facilitation after supportive contexts.

2. Method

2.1. Participants

Twenty-four right-handed young adults (12 females, M = 23.2 years, SD = 4.3) were paid to participate in the experiment after giving written consent. We discarded the data of three participants due to excessive blinking (n = 1) and signs of sleepiness as evidenced by high alpha activity throughout the experiment (n = 2). Therefore, the total sample included 21 participants (11 females, M = 23.4 years, SD = 4.5). None of the participants reported health problems or prior neurological disorders.

2.2. Stimuli and task

The three-hundred fifty-two sentences (176 high constraint and 176 low constraint) used in León-Cabrera et al. (2017) and created by Mestres-Missé et al. (2007) were employed (Table 1). According to the original paper, the mean cloze-probability sentence completions was 6.1% (SD = 10.3%) for the low constraint condition (LC) and 76%

Table 1
Sample set of sentence examples for each experimental condition with translations to English.

Condition	Sentence Context	Final word
HCC	El portero fue capaz de atrapar la	pelota
	The goalkepper managed to catch the	ball
HCI	El portero fue capaz de atrapar la	orilla
	The goalkepper managed to catch the	shore
LCC	Le ha regalado a su hijo una	pelota
	As a present she gave her son a	ball
LCI	Le ha regalado a su hijo una	orilla
	As a present she gave her son a	shore
NS	Helade algoroa seujohi nua	viaje (trip)

(SD = 17.7%) for the high constraint condition (HC). To construct a non-semantic condition (NS), we randomly picked 40 sentences from the LC condition and scattered the vowels within each word to make them semantically meaningless but grammatically plausible. The non-semantic condition was included to provide a control condition in which the sentence was formed by linguistic stimuli of similar complexity to the words of the HC and LC conditions, but without any semantic content. Each sentence had one congruent final word (their best completion) and one incongruent and implausible final word. Given the features of the congruent final words (mean word length = 6.6 letters [SD = 1.87], mean number of syllables = 2.9 [SD = 0.82] and mean word frequency = 35.18 [SD = 75.18]), we carefully selected the incongruent final words from the ESPAL database (Duchon et al., 2013) so that they matched the congruent words (mean word length = 6.6 letters [SD = 1.87], mean number of syllables = 2.88 [SD = 0.88] and mean word frequency = 33.68 [SD = 69.34]). Also matching the features of the congruent words, another intact 40 final words were chosen for the non-semantic condition (mean word length = 6.72 letters [SD = 1.77], mean number of syllables = 2.93 [SD = 0.76] and mean word frequency = 33.42 [SD = 61.74]).

We controlled potential confounds of the effect of contextual constraint in the N400 component. To do that, final words always had the same congruency status (congruent or incongruent), but whether they followed HC or LC contexts was counterbalanced across participants. For instance, half of the participants read "plane" as a congruent continuation for the HC context "I have never flown on a …" and the other half as a congruent ending for the LC context "The dot in the sky must be a …". Likewise, the word "drink" served as an incongruent word for the same sentences. Furthermore, every congruent and incongruent pair (i.e. "plane" and "drink") were matched in word length, word frequency, familiarity, imaginability and concreteness. In the case of the NS condition, the scattered sentence ended with a different word (e.g. "trip") that was matched in the same variables with the corresponding congruent and incongruent words (e.g. "plane" and "drink") in the semantic conditions.

Trials proceeded as follows (Fig. 1). First, a fixation point (a cross) appeared at the center of the screen for a period of between 1350 and 1750 ms (uniform distribution with a 50 ms step). Then, the sentence appeared, presented one word at a time for 200 ms (500 ms stimulus onset asynchrony, SOA). The font used was Courier New, with a size of 36 points. The color of the letters was black, while the background was white. A 1 s delay took place after the offset of the penultimate word, to analyse the anticipatory period of the final word. After the end of the final word, they waited for 800 ms until the blinking signal appeared (a depiction of an eye at the center of the screen). The blinking signal remained for 2 s before the next trial began.

To ensure that participants were reading attentively, a memory recognition test followed each block. In this test, ten words were visually presented one at a time; half were old words and half were new ones. Participants had to respond, by pressing a key, whether they had read that word in the previous block or not. Every word remained on the screen until an answer was supplied, followed by a 600 ms fixation cross before the presentation of the next word. The correspondence between the keys ("z" and "m") and the responses ("yes" and "no") were counterbalanced across participants. At the end of the memory task, the main task continued right away, except for even-numbered blocks, which were followed by a pause that could be resumed anytime.

2.3. Procedure

After general instructions and preparation, participants were comfortably seated in approximately 70 cm away from the computer screen that would be later used. After this, the EEG cap was set up and the state of each electrode was checked. Participants were given instructions about how to reduce artifacts by minimizing movement and to wait for a visual signal at the end of each trial to blink. After completing another unrelated task, the briefing was carried out. Participants were told to read each sentence carefully and that after each block they would have to complete a recognition test related to the final words presented during those trials.

2.4. EEG recording

Electrophysiological data (EEG; sampling rate = 500 Hz; on-line bandpass filter = .015–1000 Hz) was recorded from 29 tin scalp electrodes at standard 10/20 system positions (electrode positions: FPz, FP1/2, Fz, F3/4, F7/8, FCz, FC3/4, Cz, C3/4, CPz, CP3/4, Pz, P3/4, TP7/8, T3/4, T5/6, Oz, O1/2, left and right mastoids). The EEG signal was re-referenced offline to the mean activity of the mastoid electrodes. Vertical and horizontal electro-oculograms were recorded and used for artifact rejection. All electrode impedances were kept below 5 kΩ. Before performing statistical analysis, the data were filtered offline at 50 Hz and 60 Hz with a notch filter (to attenuate electrical line noise) and at 30Hz using a low-pass Butterworth filter (roll-off of 12 dB/oct) as implemented in ERPLAB toolbox V6.1.3 (López-Calderón and Luck, 2014). To perform artifact rejection, we excluded the epochs in which the peak-to-peak amplitude in the electro-oculograms exceeded ±85 μV (moving window = 200 ms, moving step = 20 ms) or in which activity was ±200 μV in any other channel. After this, additional visual inspection of the resulting signal was carried out for each subject individually. Using these criteria, a mean of 19.05% of trials were rejected (SD = 11.7).

2.5. ERP data analysis

The data analyses were divided into three parts that focused separately on the electrophysiological activity, 1) over the course of the sentence context (sentence-level interval), 2) in the anticipatory period of the final word (pre-word interval), and 3) at final word processing (post-word interval). We applied repeated measures ANOVA to perform confirmatory analysis on the effects for which we had a prior hypothesis

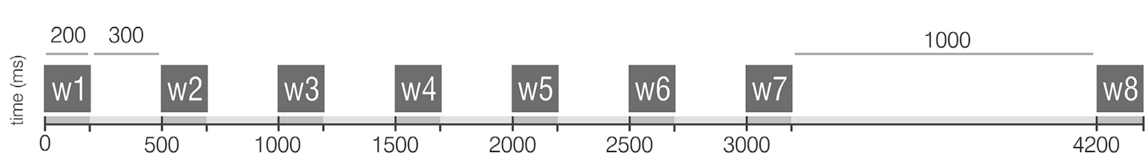

Fig. 1. Depiction of the structure of a trial. Sentences were presented one word at a time on a screen. Each word was displayed for 200 ms. A 300 ms inter-word interval was inserted between the words in the sentence context (from w1 to w7) and a 1200 ms SOA (i.e. pre-word interval) separated the penultimate (w7) and the final word (w8).

based on previous studies, and cluster-based permutation to carry out exploratory analysis and/or to adequately control for Type I errors when multiple comparisons were involved.

2.5.1. Sentence-level interval

Sentence processing was examined within an epoch length of 4200 ms that comprised the activity from the onset of the first word (w1) to the end of the delay period before the eighth and final word (w8). The epoch was baseline-corrected to its preceding 100 ms. For this analysis, we had a hypothesis regarding the direction of the differences between conditions based on a previous study (León-Cabrera et al., 2017) that showed more negative amplitudes to HC. However, as we had no strong prediction about the spatiotemporal locus of the effect in the time-course of the whole sentence, we adopted an exploratory approach and applied a non-parametrical cluster-based permutation test (Maris and Oostenveld, 2007) to analyse experimental differences in this period. This method controls the probability of a false positive (Type I error rate) in the set of multiple comparisons. Following this procedure, every sample (Channel x Time) was compared between two conditions (HC vs. LC, HC vs. NS and LC vs. NS, separately) by means of *t*-statistics. Then, an algorithm clustered the adjacent spatio-temporal samples with similar differences based on a *t*-value threshold of ± 1.72 (alpha level of 0.05 with 20 degrees of freedom, for one-tailed testing). A cluster-level statistic (permutation *p*-value) was computed under a permutation distribution of the cluster with the largest sum of t-values. The permutation distribution was approximated by a Monte Carlo method involving 15000 randomizations of the data of the two experimental conditions. Only clusters with a permutation *p*-value below 5% (critical alpha level) were considered significant. We chose one-tailed testing because –as previously stated– we had an a priori hypothesis that the amplitude would be more negative for HC compared to all the other experimental conditions (HC < LC < NS). Before running the test, we applied a low-pass filter at 5 Hz to focus solely on slow brain potentials (Brunia et al., 2011). Importantly, the cluster-based permutation procedure deals effectively with the increased probability of noise autocorrelation involved in setting such a small low-pass filter cut-off (Piai et al., 2015) and therefore it is a very adequate statistical solution for these situations.

2.5.2. Pre-word interval

Based on previous studies (Morís et al., 2013; León-Cabrera et al., 2017), we expected differences between the experimental conditions to be maximal at fronto-central sensors in the period immediately preceding the final word. We analysed the 1000 ms delay period preceding the final word (w8) in a separate window analysis, time-locked to the onset of the penultimate word (w7) and using the 100 ms pre-stimulus as a baseline (Van Petten and Kutas, 1991; King and Kutas, 1995). The differences were assessed by pre-planned repeated measures analysis of variance (ANOVA) of three time-windows in the final 600 ms of the delay period (−600 to −400 ms, −400 to −200 ms, and −200 to 0 ms) (Morís et al., 2013) on a subset of six fronto-central sites.

2.5.3. Post-word interval

The analysis of the activity elicited by the presentation of the final word focused on the N400 component. The time window for the analysis comprised the 300–500 ms time-locked to the onset of the final word (w8), with a 100 ms pre-stimulus baseline (León-Cabrera et al., 2017). The CPz electrode data was analysed using a repeated-measures ANOVA. CPz was chosen as the representative electrode based on the topographical distribution of the strongest N400 effect in the sample, although, for completeness, an analysis using a cluster of centro-parietal electrodes (Cz, CPz and Pz) was also carried out (Supplementary Materials). We expected to find an earlier onset of the N400 congruency effect for HC compared to LC (León-Cabrera et al., 2017). Given that we had an a priori hypothesis about the location of the effect, we ran a single-sensor cluster-based permutation test (Maris and Oostenweld, 2007; Bullmore et al., 1999) on the CPz electrode to determine the temporal onset of the effect for each level of semantic constraint (HCI minus HCC, and LCI minus LCC, separately) in the 650 ms after final word onset. Again, we chose one-tailed testing because we predicted that incongruent words would have more negative mean amplitudes than congruent words. We did not plan to include the NS condition in these analyses because its functional interpretation is unclear in the post-word interval, we nevertheless added it in a grand-averaged ERP in the Supplementary Materials for those interested.

Statistical analyses were run with the FieldTrip toolbox (version 28-01-2018) (Oostenveld et al., 2011) and with SPSS 21. In repeated-measures ANOVA, the Greenhouse-Geisser correction was applied whenever the sphericity assumption was not met. Corrected *p*-values are reported as p_{GG}. In case a theoretically relevant interaction was found to be significant, we disentangled the effect by means of post-hoc *t* tests.

3. Results

3.1. Sentence-level slow sustained negativity

The grand average ERPs of the sentence context (Fig. 2) revealed widespread slow brain potentials with amplitude differences across experimental conditions that were confirmed statistically by the non-parametric cluster-based permutation test (Fig. 3).

A significant negative cluster was found both between HC and LC ($p < .001$) and HC and NS ($p < .001$), consistent with the hypothesis that voltage amplitudes would be more negative for higher levels of semantic constraint. In the HC and LC contrast, the cluster started around the fifth word (w5; from 2200 ms onwards), whereas it started earlier in the HC and NS comparison, around the fourth word (w4; from 1600 ms onwards). In both cases, the cluster prolonged to the end of the sentential context and appeared to be maximal at fronto-central sites. No significant clusters were found between LC and NS. As a measure of the robustness of these effects, we replicated the finding of these significant negative clusters using a more conservative testing (Supplementary Materials).

3.2. Local slow negative potential

We investigated the delay preceding the presentation of the final word, as we expected brain activity to portray the maximal differences in the slow preceding negativity during this period (Fig. 4).

A repeated-measures ANOVA was conducted on the mean amplitudes of six fronto-central electrodes (F3, Fz, F4, FC3, FCz and FC4), involving four factors: Condition (3 levels, HC, LC, NS), Time (3 levels, 400–600 ms, 600–800 ms, and 800–1000 ms), Electrode (2 levels, frontal, fronto-central), and Laterality (3 levels, left, central, right). Given the computation of multiple comparisons, Bonferroni-corrected post-hoc *t* tests were carried out to further explore the effects found. The corrected *p*-values are indexed as p_{BF}.

Overall mean amplitude differences between each level of semantic constraint were quantified as a main effect of Condition ($F(2,40) = 19.08$, $p < .001$, $\eta_p^2 = 0.48$). Equally to what we observed in the whole sentence analysis, the mean amplitude linearly changed with the level of semantic constraint ($F(1,20) = 70.75$, $p < .001$, $\eta_p^2 = 0.78$), with HC being significantly more negative than LC ($p_{BF} = .031$) and NS ($p_{BF} < .001$), and LC being comparatively more negative than NS ($p_{BF} = .036$).

Consistent with the typical evolution of the slow stimulus preceding negativity, there was a significant main effect of Time ($F(2,40) = 17.05$, $p_{GG} < .001$, $\eta_p^2 = 0.46$) that stemmed from a steeper negative shift in the mean amplitude of the final 200 ms, as can be seen in the scalp maps (Fig. 5). More specifically, *t* tests showed that the mean amplitudes between the first and second time-windows did not differ, but the third time-window (corresponding to the last 200 ms) became significantly

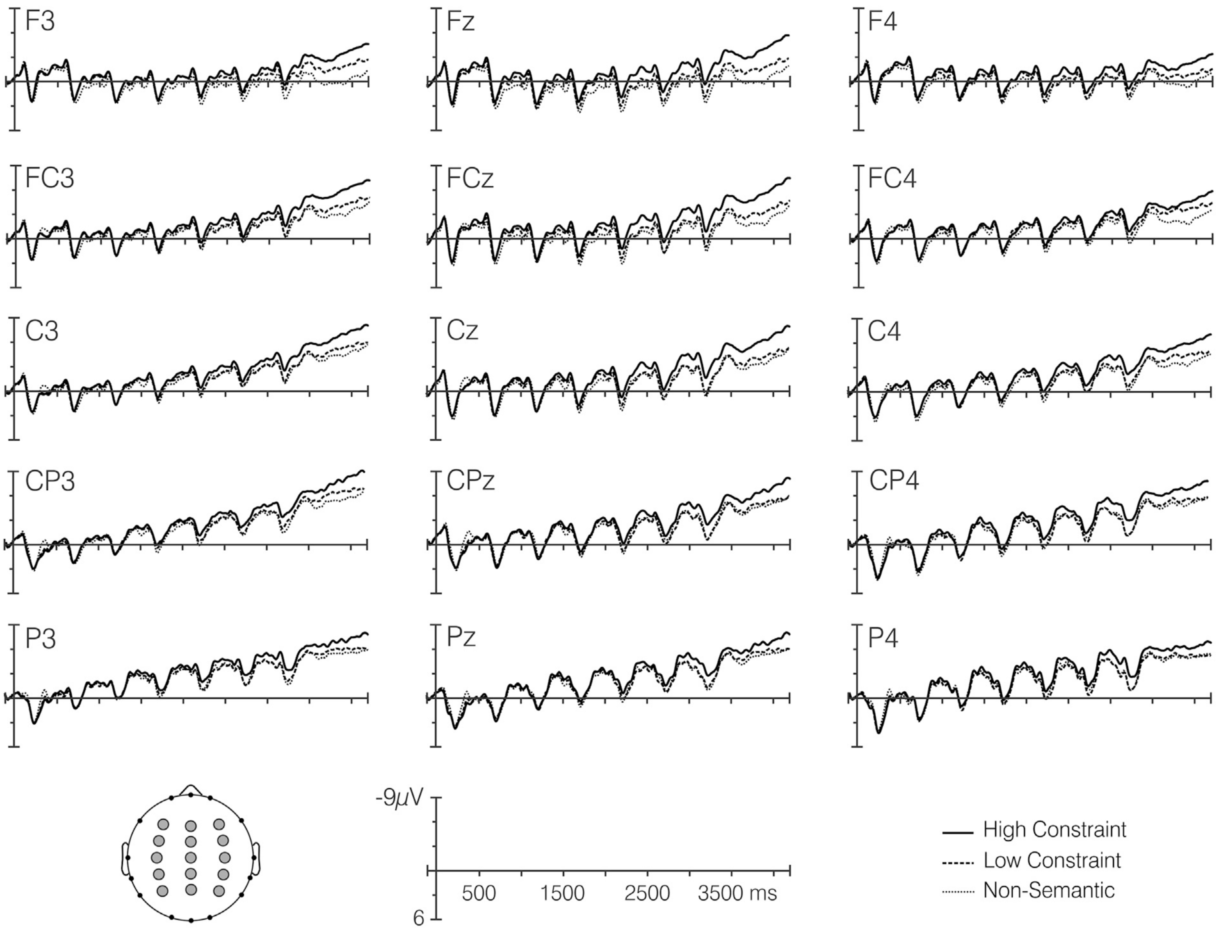

Fig. 2. Grand-averaged ERPs across sentence contexts (0–3200 ms) and the following final pre-word anticipatory period (3200–4200 ms) showing the sentence-level slow cortical modulations at frontal (F3, Fz, F4), fronto-central (FC3, FCz, FC4), central (C3, Cz, C4), centro-parietal (CP3, CPz, CP4) and parietal (P3, Pz, P4) sites in the three experimental conditions (High Constraint, HC; Low Constraint, LC, and Non-Semantic, NS). Negative is plotted upward.

more negative (p_{BF} <.001 in both contrasts). Indeed, the Condition × Time interaction was significant ($F(2,80) = 3.09$, $p = .02$, $\eta_p^2 = 0.13$). In the earliest time-window, from −600 to −400 ms, HC and LC did not differ, whereas both were more negative than NS (HC vs. LC, $p = .242$, HC vs. NS, $p < .001$, LC vs. NS, $p = .036$). Later, from −400 to −200 ms, all the conditions differed significantly between them. Again, as observed in the main effect of condition, showing more negative amplitudes in HC compared to LC ($p_{BF} = .009$) and NS (p_{BF} <.001), and LC being more negative than NS ($p_{BF} = .004$). Eventually, in the final 200 ms, HC remained more negative than LC ($p_{BF} = .002$) and NS (p_{BF} <.001), but the difference between LC and NS disappeared.

Finally, the significant Condition × Laterality interaction ($F(2,80) = 2.74$, $p = .03$, $\eta_p^2 = 0.12$) confirmed the prediction that the effect of semantic constraint was most pronounced at left sites. In fact, at right and central positions only HC and NS diverged ($p_{BF} = .02$), whereas all the conditions were different at left sites (all p_{BF} <.05).

There were also main effects of Electrode (p_{BF} <.001) and Laterality ($p_{BF} = .002$), and significant interactions of Position x Laterality, Time x Electrode, Time x Laterality, and Time x Position x Laterality (all p_{BF} <.05). We do not expand on these results because they do not interact with the effect of the experimental manipulation and thus are considered uninformative in this context.

3.3. N400 component

We conducted a repeated-measures ANOVA on the CPz electrode data using the time window of 300–500 ms from the onset of the final word to investigate the N400 component (Fig. 6A), using two factors, Contextual constraint (2 levels, high, low) and Congruency (2 levels, congruent, incongruent). The analysis showed significant main effects of Contextual constraint, ($F(1,20) = 30.9$, $p < .001$, $\eta_p^2 = 0.60$), and Congruency ($F(1,20) = 30.32$, $p < .001$, $\eta_p^2 = 0.60$), but no significant interaction between the factors. Post-hoc *t* tests revealed the classical N400 effect, whereby incongruent words produced more negative amplitudes than congruent words, both within high (mean amplitudes, HCI = 2.65 μV, HCC = 6.53 μV, $p < .001$) and low (mean amplitudes, LCI = −0.16 μV, LCC = 2.19 μV, $p = .002$) constraint conditions. To test the robustness of the effect, we computed an equivalent ANOVA using a centro-parietal cluster (Cz, CPz and Pz) that replicated the same results (Supplementary Materials).

Single-sensor cluster-based permutation tests on the difference waveforms of the constraint conditions (HCI minus HCC, and LCI minus LCC) at CPz, confirmed the prediction that congruency effects started earlier for high than low levels of semantic constraint (Fig. 6B. Left). A significant cluster was found from 216 to 522 ms in the high constraint contrast and from 302 to 564 ms in the low constraint condition (all cluster *p*-values < .001, one-tailed testing). Lastly, we report the time segment in which *t* values of the comparison between the two difference waveforms exceed the critical value under a one-tailed *t*-distribution (±1.72, for one-tailed testing). The period comprises a window from 278 to 414 ms (136 ms duration) (Fig. 6B. Right). The topographical representation of the difference waveforms visibly exhibits the distinct temporal evolution of the effect (Fig. 6C).

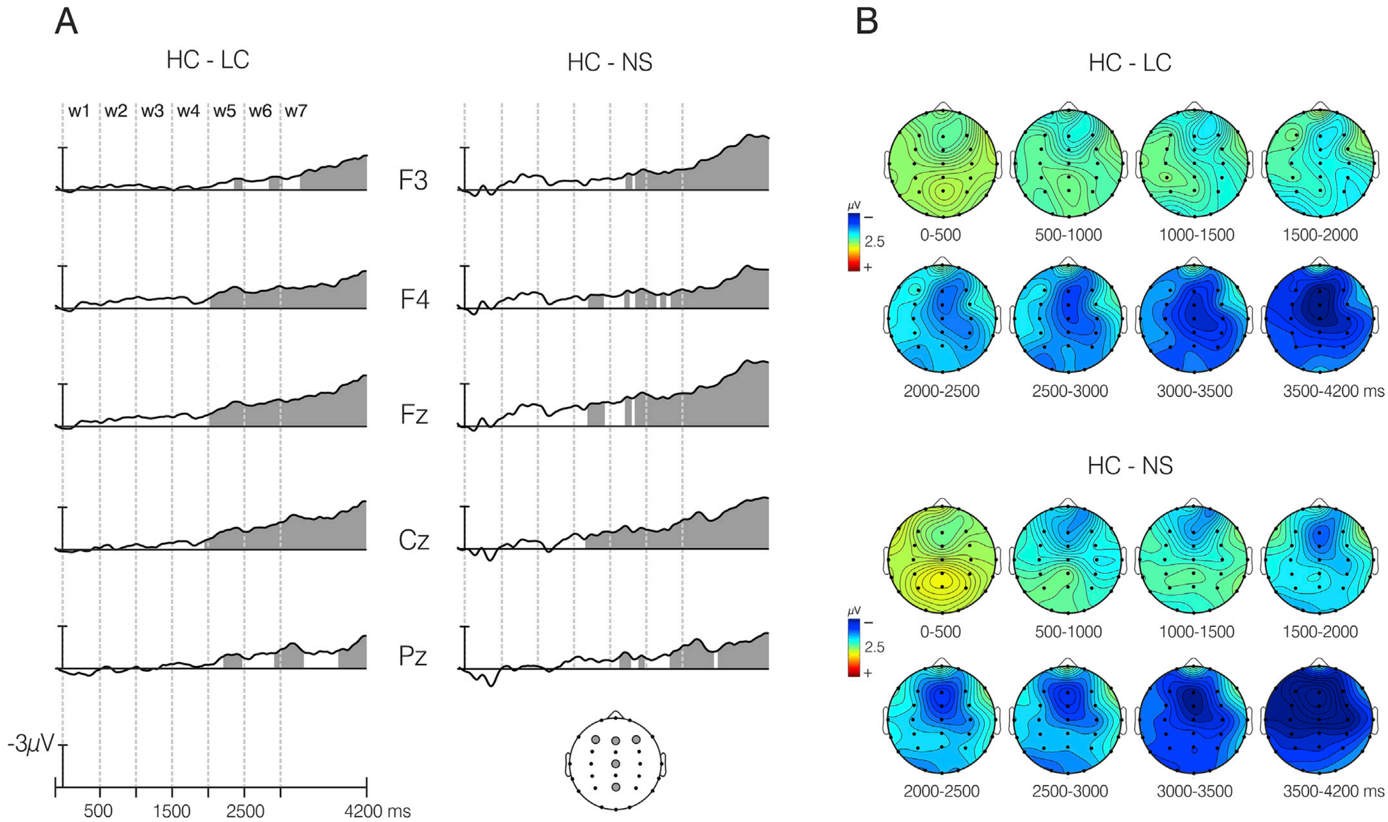

Fig. 3. A: Difference waveforms of the statistically significant contrasts (HC minus LC and HC minus NS) in the non-parametrical cluster-based permutation analysis that comprised the full epoch (0–4200 ms) time-locked to the onset of the sentence context, at 5 representative electrodes (F3, F4, Fz, Cz and Pz). The onset of each word is marked with a vertical dotted line and the corresponding word position (w_x). The grey portions represent the statistically significant areas of the cluster. Negative is plotted upward. B: Scalp maps displaying the temporal evolution of the sentence-level sustained negativity based on mean amplitude differences (HC minus LC and HC minus NS) in eight consecutive epochs of 500 ms. The range of voltage values for the maps is ±2.5 μV.

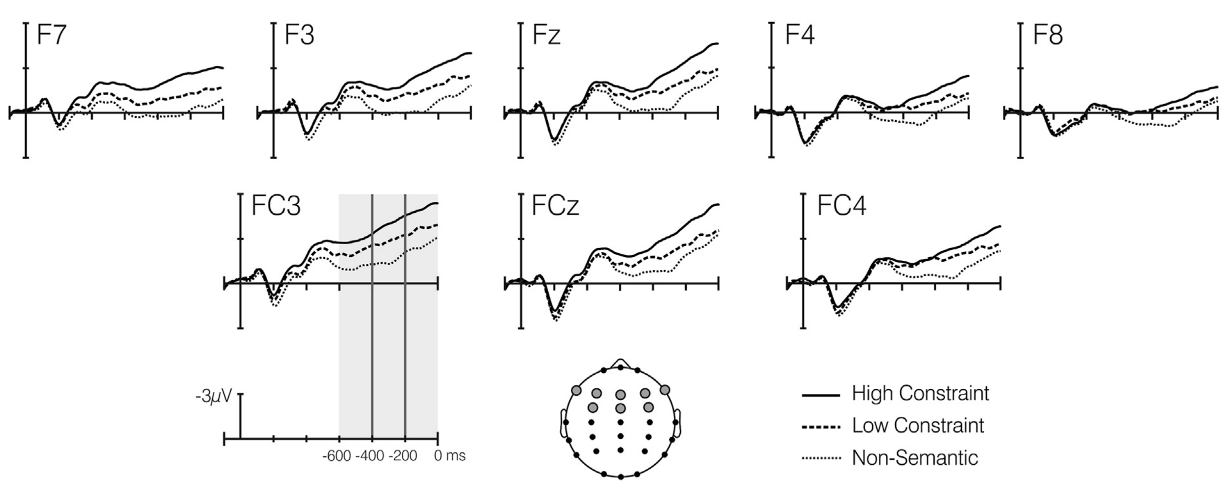

Fig. 4. Grand-averaged ERPs to the pre-word interval showing the local slow negative potential preceding the presentation of the final word at frontal (Fz, F3/4, F7/8) and fronto-central (FCz, FC3/4) sites, where the differences were maximal. The ERPs are time-locked to the onset of the penultimate word. The grey area indicates the time intervals that were separately subjected to statistical analysis. The time values (X axis) are referenced to the onset of the final word.

4. Discussion

This study investigated sustained ERP correlates associated to the build-up of predictions during sentence processing. To address this, participants read, word by word, sentences with different degrees of semantic constraint (high, low or none) and a constant 1 s interval between the penultimate and the final word (congruent or incongruent). In support of our main hypothesis, processing differences as a function of semantic constraint were captured by broadly distributed sustained negativities that emerged early and gradually over time, with increasingly larger amplitudes for high than low levels of constraint. The differences were maximal in the interval immediately preceding the closing word, which elicited a left-dominant local slow negative potential with a graded amplitude modulation to the level of constraint, replicating a previous report in speech comprehension (León-Cabrera et al., 2017). The presentation of the closing word elicited a canonical modulation of

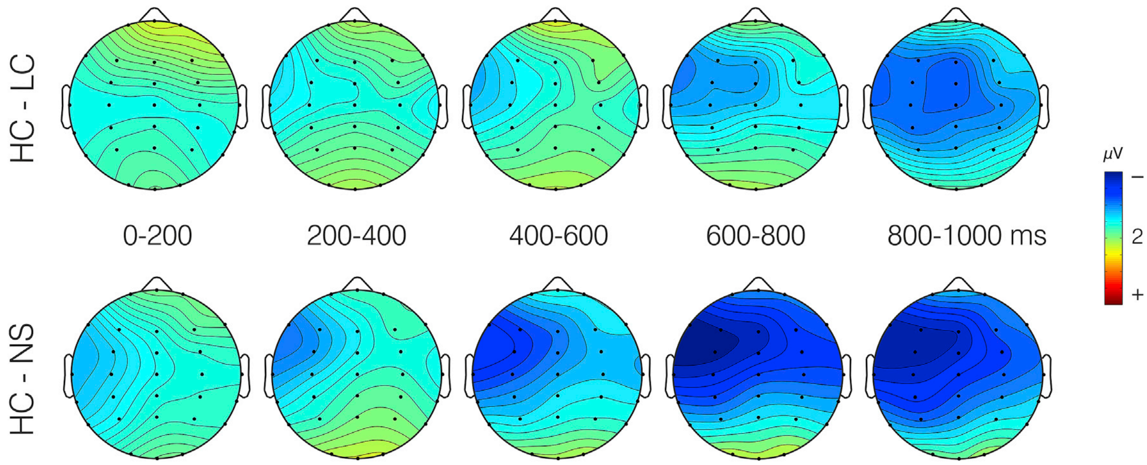

Fig. 5. Scalp maps of the mean amplitudes of the difference waveforms (HC minus LC and HC minus NS) showing the temporal evolution of the local slow negative potential that developed in the 1000 ms pre-word interval in five consecutive time-windows of 200 ms. The range of voltage values for the maps is ±2 µV.

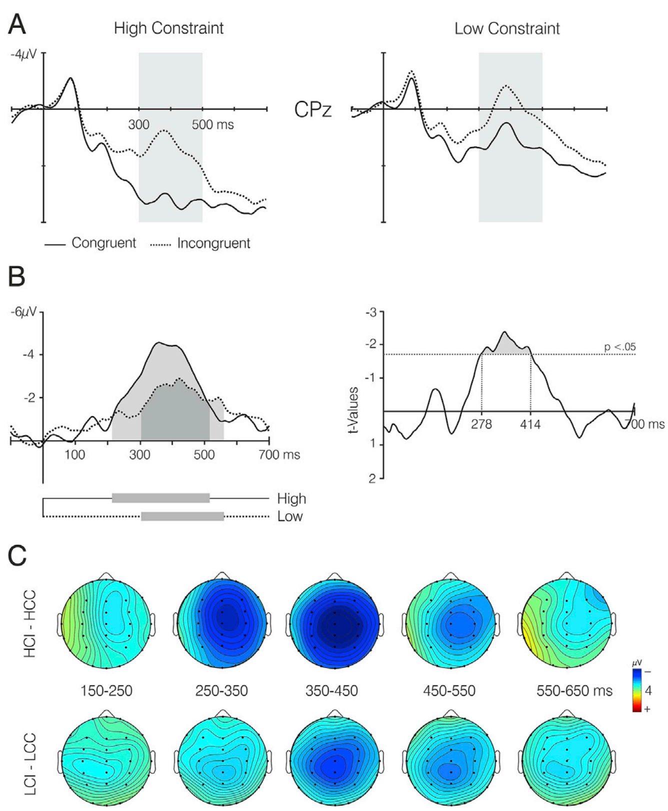

Fig. 6. A: Grand-averaged ERPs for each condition showing the N400 component at the CPz electrode, time-locked to the onset of the final word. The grey area indicates the interval where the amplitudes were statistically compared. B. Left: Difference waveforms (Incongruent minus Congruent) displaying the N400 effect to each condition of semantic constraint. The grey segments correspond to the statistically significant clusters. B. Right: Difference waveforms of t-value evolution of the differences between the two waveforms. The grey dotted line indicates the point at which the waves start to differ. C. Scalp maps of the mean amplitude difference between the difference waveforms in Figure 6B every 100 ms in the 150–650 ms interval, showing the earlier onset of the effect for the HC compared to the LC condition. The voltage values for the maps ranged ± 4 µV.

the N400 component to semantic fit and an earlier onset of the N400 effect for high compared to low contextual constraint. As we argue henceforth, the features of the local and sentence-level negativities support their functional interpretation as indices of predictive processing and the N400 effects are also in line with this interpretation. We focus first on the local slow potential, given the availability of similar instances in the literature, then turn to the novel contributions of the sentence-level modulation and, finally, to the N400 effects.

The local slow potential in the pre-word interval replicates the finding of León-Cabrera et al. (2017) using a similar task in speech comprehension. In both studies, we observed the development of a slow negative potential over fronto-central sites and a graded pattern of amplitude differences as a function of contextual constraint, with more negative amplitudes for stronger levels of constraint (HC < LC < NS). We noted that the observed potential shared many features with the SPN – a frontal locus, a gradual amplitude increase over time, and a modulation that agrees with the predictability of the upcoming event (Walter, 1964; Brunia and Damen, 1988). The fact that we found a very similar slow potential using different input modalities (i.e. auditory and visual) bolsters the idea that it is associated to a top-down process, such as semantic anticipation. If it reflected unspecific anticipatory attention towards an imperative stimulus (in this case, the final word), one would expect different topographical distributions for each sensory modality (Brunia and Van Boxtel, 2004).

The left-dominant topographical distribution of the amplitude differences in the local slow potential is reminiscent of prior reports that are thought to reflect the manipulation of verbal representations. For instance, Heil et al. (1996) had participants learn the associations between a drawing and one, two or three words. During recall, a frontal negativity developed prior to the presentation of the word. Resembling the current findings, the slow wave was maximal at left anterior areas (in their case, the F3 electrode), where the amplitude was monotonically related to the number of verbal representations that had to be retrieved (see also Rösler et al., 1995). Other studies have associated left anterior negativities to the elaborative encoding of verbal associations (Lang et al., 1988), the retention of verbal material (Ruchkin et al., 1990) or the active maintenance of verbal items (Khader et al., 2007). More generally, in studies of language comprehension using more elaborate materials, transient left anterior negativities have been typically associated to WM operations such as the temporary storage or the retrieval of specific elements (Kluender and Kutas, 1993; Fiebach et al., 2002; Piai et al., 2013; Matzke et al., 2002).

The possibility that the amplitude modulation of the local slow potential is associated to the retrieval and/or maintenance of verbal information fits nicely with a prediction-based interpretation of the effect – the sentence context could have led to the retrieval of semantic features associated to the most probable continuations before their bottom-up input (Kutas et al., 2011, for a review). If the context is sufficiently constraining, as in the HC condition, even specific lexical items could have been pre-activated (Boudewyn et al., 2015), but this possibility cannot be tested with the current task. Following this, contextually-driven semantic and/or lexical predictions would establish different levels of expectancy towards the upcoming word. In a similar token, Grisoni et al. (2017) attributed to semantic prediction their finding of a slow negative potential that was maximal over the motor areas conceptually-related to the final word. The distinct topographical distribution may be accounted by task differences, such as the use of action verbs (instead of nouns, in our case) or the restriction of candidates to only two possible semantic categories (instead of an unbounded set of categories). This could have led to the activation of qualitatively distinct information, thus tapping on distinct cortical regions. Nevertheless, both studies converge in finding anticipatory slow potentials that are modulated by semantic constraint in a way that is consistent with contextually-driven anticipation.

Importantly, the present study reveals that the local slow potential may be the continuation of slower, sentence-level process. Based on interactive models of language comprehension, we hypothesized that contextual constraint would have an earlier and sustained impact on brain activity during sentence processing. Accordingly, we found that HC contexts diverged already at the fourth word from NS contexts, and at the fifth word from LC contexts. These differences were captured by a slow sustained negativity that developed gradually across sentence processing and that was broadly distributed over frontal and central areas. On the one hand, the sustained negativity could result from the sum of successive word-level integration processes, given that it spanned over several words. Under this view, it could reflect the incremental construction of a meaning representation. Accordingly, the pattern of differences could be explained by the relative ease or difficulty integrating the words that each type of context affords; the greater semantic information provided by HC contexts, relative to LC and NS contexts, would have facilitated the incorporation of the words into the higher-level representation of the sentence, and possibly also the achievement of a clearer message-level interpretation. On the other hand, the sentence-level modulation could be capturing more than word-by-word context updating. The differences are also consistent with more recent prediction-based models of language comprehension, whereby contextual information influences the state of the language processing system before bottom-up input is received. In fact, the early and progressive evolution of the differences agrees well with the notion that predictions about upcoming concepts (and/or some of their features) are computed and continuously updated online (for a review, Kuperberg and Jaeger, 2016).

As we described in the Introduction, high-level integration during sentence processing has been associated with slow positive (rather than negative) shifts (see Kutas and King, 1996, for a review). For instance, Van Petten and Kutas (1991) found that reading semantically meaningful and congruent sentences lead to a word-by-word decrement of the N400 component, resulting in a slow positive shift. Cross-clause negativities have been linked to increased WM demands when processing sentences that have more difficult syntactic configurations (e.g., King and Kutas, 1995; Müller et al., 1997; Fiebach et al., 2002) or in which the linguistic order of presentation of the events does not match the conceptual order (Münte et al., 1998). Following this, the polarity of the sentence-level modulation could be indicative of differences in memory demands as a function of contextual constraint. For instance, some authors have suggested that memory demands could be involved in maintaining multiple representations active in parallel when the context allows for several interpretations (Wlotko and Federmeier, 2012). Interestingly, the theoretical models that were used account for the effects in some aforementioned studies (King and Kutas, 1995; Müller et al., 1997; Fiebach et al., 2002) purport that processing, and memory costs are incurred by the temporary maintenance of partial linguistic information WM (Clifton and Frazier, 1989; Gibson, 1998). Although these models were devised to account specifically for syntactic parsing phenomena (i.e. filler-gap dependencies), similar cognitive mechanisms could be involved during semantic processing in some situations. If predictive processing is engaged, memory demands could be involved in maintaining predicted representations available within the cognitive system, at least until bottom-up input is received. In this line, some accounts of prediction have suggested that items are predictively added to the contextual representation that is being held in WM (Lau et al., 2013), thus implying a potential relationship between working memory and predictive processes in language.

The LC and the NS condition did not differ significantly in the sentence-interval. This may seem surprising, given that in the NS condition it was impossible for participants to build a meaningful representation, whereas they should have been able to do so in the LC condition. However, it should be noted that the LC condition in this task had a very low semantic constraint – the mean cloze probability was below 10%. Therefore, it is possible that participants could not accrue much information online and were unable to either build a robust contextual representation or generate online predictions in the LC condition. However, the two conditions differed significantly in the pre-word interval, suggesting that the LC condition could have eventually offered some preparatory advantage over the NS condition. At this point, we should also consider that some features of the task – such as the cloze-like structure of the sentences (Ferreira and Lowder, 2016), word-by-word presentation of stimuli or/and the introduction of a delay – may have encouraged participants to adopt a more strategic approach for sentence comprehension that would not be engaged in other situations (Brothers et al., 2017). For example, if the task structure had been less favorable for prediction, or sentence comprehension had assisted an ulterior task goal, perhaps the qualitative distinction between NS and LC sentences would have been more relevant and the two conditions would have differed substantially also during sentence processing.

Finally, we examined processing differences once the final word was presented. We found the classical modulation of the amplitude of the N400 component to contextual constraint (e.g. Kutas and Hilyard, 1984) and to the congruency of the sentence completions (e.g. Kutas and Hillyard, 1984; Kutas and Federmeier, 2000). In addition, we found an earlier onset of the N400 effect (incongruent minus congruent) for HC compared to LC contexts, which replicates a previous observation using auditory stimuli (León-Cabrera et al., 2017). On the one hand, the latency effect could stem from facilitated processing for congruent words in the HC contexts (Federmeier and Kutas, 1999). This is consistent with the functional interpretation of the local and sentence-level slow negativities that precede word processing as correlates of prediction, whereby semantic features associated to the best completion would have become activated, consequently allowing for faster word integration upon receipt. On the other hand, in case stronger forms of prediction took place, HC contexts could have led to the lexical pre-activation of the best completion specifically, allowing to detect a mismatch sooner through straightforward comparison with the bottom-up input (Boudewyn et al., 2015). However, in previous studies, mismatch effects have been usually attributed to even earlier responses, such as an N200 (Boudewyn et al., 2015) or an N250 (Cermolacce et al., 2014). In these studies, the validity of the predictions was almost 100% in the HC condition, either because contexts were maximally constraining (Boudewyn et al., 2015) or because fixed expressions (i.e. proverbs) were used. This could have encouraged stronger forms of prediction (i.e. lexical pre-activation) than in this task, where predictions were disconfirmed by incongruent endings in half of the sentences. Although the present data is insufficient to tell apart the exact underlying mechanism, the finding of a latency effect supports the involvement of predictions at the semantic level during comprehension in this task.

Altogether, these findings contribute significantly to the available literature by providing clear evidence of early, continuous and gradual impact of top-down information on sentence comprehension. The sentence-level negativity represents one of the few instances of slow and sustained changes associated to semantic constraint, extending previous studies that focused on more transient effects (León-Cabrera et al., 2017;

Grisoni et al., 2017). Overall, the pattern of results adds to previous evidence in support of prediction in language comprehension, in the sense that the state of the system changes prior to bottom-up input (in this case, at least prior to the final word of the sentence) as a function of top-down information. Interestingly, the local slow potential could be reflecting retrieval or maintenance of a verbal representation at some representational level, but we did not find clear evidence of the pre-activation of specific lexical representation either in the pre- or at the post-word intervals. Future research is needed to investigate the potential relationship between the observed slow negativities and stronger forms of prediction, as well as differences in memory demands. The current experimental approach seems to be a valuable method to reveal potential markers of predictive processes engaged during sentence comprehension, and future studies should address whether these results can be transferred to different experimental set-ups or to more naturalistic comprehension contexts.

Conflicts of interest

The authors declare no disclosure of financial interests and potential conflict of interest.

Acknowledgments

This research was supported by grant PSI2013-43516-R of the Spanish Ministry of Economy and Competitiveness. PLC was funded with a pre-doctoral grant FPU15/05554 (FPU "Ayudas para la Formación de Profesorado Universitario") of the Spanish Ministry of Education, Culture and Sport. None of these funding sources had any involvement in the conduct of the research or the preparation of the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.01.005.

References

Boudewyn, M.A., Long, D.L., Swaab, T.Y., 2015. Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension. Cognit. Affect Behav. Neurosci. 15 (3), 607–624. https://doi.org/10.3758/s13415-015-0340-0.

Brothers, T., Swaab, T.Y., Traxler, M.J., 2017. Goals and strategies influence lexical prediction during sentence comprehension. J. Mem. Lang. 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002.

Brunia, C.H.M., Damen, E.J.P., 1988. Distribution of slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. Electroencephalogr. Clin. Neurophysiol. 69 (3), 234–243.

Brunia, C.H.M., Van Boxtel, G.J.M., 2004. Anticipatory attention to verbal and non-verbal stimuli is reflected in a modality-specific SPN. Exp. Brain Res. 156 (2), 231–239. https://doi.org/10.1093/oxfordhb/9780195374148.013.0108.

Brunia, C.H., van Boxtel, G.J., Bocker, K.B., 2011. The Oxford handbook of event-related potential components.

Bullmore, E.T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., Brammer, M.J., 1999. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. IEEE Trans. Med. Imag. 18 (1), 32–42. https://doi.org/10.1109/42.750253.

Cermolacce, M., Scannella, S., Faugère, M., Vion-Dury, J., Besson, M., 2014. All that glitters is not… alone". Congruity effects in highly and less predictable sentence contexts. Neurophysiol. Clin. 44 (2), 189–201. https://doi.org/10.1016/j.neucli.2014.04.001.

Clifton, C., Frazier, L., 1989. Comprehending sentences with long-distance dependencies. In: Linguistic structure in language processing. Springer, Dordrecht, pp. 273–317. https://doi.org/10.1007/978-94-009-2729-2_8.

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

Dikker, S., Pylkkänen, L., 2013. Predicting language: MEG evidence for lexical preactivation. Brain Lang. 127 (1), 55–64. https://doi.org/10.1016/j.bandl.2012.08.004.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., Carreiras, M., 2013. EsPal: One-stop shopping for Spanish word properties. Behav. Res. Methods 45 (4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1.

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

Federmeier, K.D., Kutas, M., 1999. A rose by any other name: Long-term memory structure and sentence processing. J. Mem. Lang. 41 (4), 469–495. https://doi.org/10.1006/jmla.1999.2660.

Ferreira, F., Lowder, M.W., 2016. Prediction, information structure, and good-enough language processing. In: Psychology of Learning and Motivation, vol. 65. Academic Press, pp. 217–247. https://doi.org/10.1016/bs.plm.2016.04.002.

Fiebach, C.J., Schlesewsky, M., Friederici, A.D., 2002. Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. J. Mem. Lang. 47 (2), 250–272. https://doi.org/10.1016/S0749-596X(02)00004-9.

Freunberger, D., Roehm, D., 2016. Semantic prediction in language comprehension: evidence from brain potentials. Language, cognition and neuroscience 31 (9), 1193–1205. https://doi.org/10.1080/23273798.2016.1205202.

Gibson, E., 1998. Linguistic complexity: Locality of syntactic dependencies. Cognition 68, 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1.

Grisoni, L., Miller, T.M., Pulvermüller, F., 2017. Neural correlates of semantic prediction and resolution in sentence processing. J. Neurosci. 2800–2816. https://doi.org/10.1523/JNEUROSCI.2800-16.2017.

Hackley, S.A., Valle-Inclán, F., Masaki, H., Hebert, K., 2014. Stimulus-preceding negativity (SPN) and attention to rewards. Cognitive electrophysiology of attention: Signals of the mind, pp. 216–225. https://doi.org/10.1016/B978-0-12-398451-7.00017-8.

Huettig, F., Mani, N., 2016. Is prediction necessary to understand language? Probably not. Language, Cognition and Neuroscience 31 (1), 19–31. https://doi.org/10.1080/23273798.2015.1072223.

Heil, M., Rösler, F., Hennighausen, E., 1996. Topographically distinct cortical activation in episodic long-term memory: The retrieval of spatial versus verbal information. Mem. Cognit. 24 (6), 777–795. https://doi.org/10.3758/BF03201102.

Jackendoff, R., 2003. Précis of foundations of language: brain, meaning, grammar, evolution. Behav. Brain Sci. 26 (6), 651–665. https://doi.org/10.1017/S0140525X03000153.

Jackendoff, R., 2007. A parallel architecture perspective on language processing. Brain Res. 1146, 2–22. https://doi.org/10.1016/j.brainres.2006.08.111.

Khader, P., Ranganath, C., Seemüller, A., Rösler, F., 2007. Working memory maintenance contributes to long-term memory formation: evidence from slow event-related brain potentials. Cognit. Affect Behav. Neurosci. 7 (3), 212–224. https://doi.org/10.3758/CABN.7.3.212.

King, J.W., Kutas, M., 1995. Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. J. Cognit. Neurosci. 7 (3), 376–395. https://doi.org/10.1162/jocn.1995.7.3.376.

Kluender, R., Kutas, M., 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. J. Cognit. Neurosci. 5 (2), 196–214. https://doi.org/10.1162/jocn.1993.5.2.196.

Kuperberg, G.R., Jaeger, T.F., 2016. What do we mean by prediction in language comprehension? Language, cognition and neuroscience 31 (1), 32–59. https://doi.org/10.1080/23273798.2015.1102299.

Kutas, M., 1997. Views on how the electrical activity that the brain generates reflects the functions of different language structures. Psychophysiology 34 (4), 383–398. https://doi.org/10.1111/j.1469-8986.1997.tb02382.x.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cognit. Sci. 4 (12), 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). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

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

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

Kutas, M., King, J.W., 1996. The potentials for basic sentence processing: Differentiating integrative processes. In: Ikeda, I., McClelland, J.L. (Eds.), Attention and performance XVI. MIT Press.

Kutas, M., DeLong, K.A., Smith, N.J., 2011. A look around at what lies ahead: Prediction and predictability in language processing. Predictions in the brain: Using our past to generate a future, p. 190207.

Kutas, M., Van Petten, C., Besson, M., 1988. Event-related potential asymmetries during the reading of sentences. Electroencephalogr. Clin. Neurophysiol. 69 (3), 218–233. https://doi.org/10.1016/0013-4694(88)90131-9.

Lang, W., Lang, M., Podreka, I., Steiner, M., Uhl, F., Suess, E., Müller, C., Deecke, L., 1988. DC-potential shifts and regional cerebral blood flow reveal frontal cortex involvement in human visuomotor learning. Exp. Brain Res. 71 (2), 353–364.

Lau, E.F., Holcomb, P.J., Kuperberg, G.R., 2013. Dissociating N400 effects of prediction from association in single-word contexts. J. Cognit. Neurosci. 25 (3), 484–502.

León-Cabrera, P., Rodríguez-Fornells, A., Morís, J., 2017. Electrophysiological correlates of semantic anticipation during speech comprehension. Neuropsychologia 99, 326–334. https://doi.org/10.1016/j.neuropsychologia.2017.02.026.

López-Calderon, J., Luck, S.J., 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front. Hum. Neurosci. 8, 213. https://doi.org/10.3389/fnhum.2014.00213.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Matzke, M., Mai, H., Nager, W., Rüsseler, J., Münte, T., 2002. The costs of freedom: an ERP–study of non-canonical sentences. Clin. Neurophysiol. 113 (6), 844–852. https://doi.org/10.1016/S1388-2457(02)00059-7.

Mestres-Missé, A., Rodriguez-Fornells, A., Münte, T.F., 2007. Watching the brain during meaning acquisition. Cerebr. Cortex 17 (8), 1858–1866. https://doi.org/10.1093/cercor/bhl094.

Morís, J., Luque, D., Rodríguez-Fornells, A., 2013. Learning-induced modulations of the stimulus-preceding negativity. Psychophysiology 50 (9), 931–939. https://doi.org/10.1111/psyp.12073.

Morris, R.K., 2006. Lexical processing and sentence context effects. In: Handbook of Psycholinguistics, second ed., pp. 377–401 https://doi.org/10.1016/B978-012369374-7/50011-0.

Müller, H.M., King, J.W., Kutas, M., 1997. Event-related potentials elicited by spoken relative clauses. Cognit. Brain Res. 5 (3), 193–203. https://doi.org/10.1016/S0926-6410(96)00070-5.

Münte, T.F., Schiltz, K., Kutas, M., 1998. When temporal terms belie conceptual order. Nature 395 (6697), 71. https://doi.org/10.1038/25731.

Nieuwland, M.S., Van Berkum, J.J., 2006. When peanuts fall in love: N400 evidence for the power of discourse. J. Cognit. Neurosci. 18 (7), 1098–1111. https://doi.org/10.1162/jocn.2006.18.7.1098.

Nieuwland, M.S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Mézière, D., 2018. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. eLife 7, e33468. https://doi.org/10.7554/eLife.33468.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 1, 2011. https://doi.org/10.1155/2011/156869.

Piai, V., Dahlslätt, K., Maris, E., 2015. Statistically comparing EEG/MEG waveforms through successive significant univariate tests: How bad can it be? Psychophysiology 52 (3), 440–443. https://doi.org/10.1111/psyp.12335.

Piai, V., Meyer, L., Schreuder, R., Bastiaansen, M.C., 2013. Sit down and read on: Working memory and long-term memory in particle-verb processing. Brain Lang. 127 (2), 296–306. https://doi.org/10.1016/j.bandl.2013.09.015.

Piai, V., Roelofs, A., Maris, E., 2014. Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. Neuropsychologia 53, 146–156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014.

Rommers, J., Dickson, D.S., Norton, J.J., Wlotko, E.W., Federmeier, K.D., 2017. Alpha and theta band dynamics related to sentential constraint and word expectancy. Language, cognition and neuroscience 32 (5), 576–589. https://doi.org/10.1080/23273798.2016.1183799.

Rösler, F., Heil, M., Hennighausen, E., 1995. Exploring memory functions by means of brain electrical topography: a review. Brain Topogr. 7 (4), 301–313. https://doi.org/10.1007/BF01195256.

Ruchkin, D.S., Johnson Jr., R., Canoune, H., Ritter, W., 1990. Short-term memory storage and retention: An event-related brain potential study. Electroencephalogr. Clin. Neurophysiol. 76 (5), 419–439. https://doi.org/10.1016/0013-4694(90)90096-3.

Van Berkum, J.J., Brown, C.M., Hagoort, P., 1999. Early referential context effects in sentence processing: Evidence from event-related brain potentials. J. Mem. Lang. 41 (2), 147–182. https://doi.org/10.1006/jmla.1999.2641.

Van Berkum, J.J., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exp. Psychol. Learn. Mem. Cognit. 31 (3), 443. https://doi.org/10.1037/0278-7393.31.3.443.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cognit. 25 (2), 394. https://doi.org/10.1037/0278-7393.25.2.394.

Van Petten, C., Kutas, M., 1991. Influences of semantic and syntactic context on open-and closed-class words. Mem. Cognit. 19 (1), 95–112. https://doi.org/10.3758/BF03198500.

Vos, S.H., Gunter, T.C., Schriefers, H., Friederici, A.D., 2001. Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. Lang. Cognit. Process. 16 (1), 65–103. https://doi.org/10.1080/01690960042000085.

Walter, W.G., 1964. Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. Nature 230, 380–384. https://doi.org/10.1038/203380a0.

Wang, L., Hagoort, P., Jensen, O., 2018. Gamma Oscillatory Activity Related to Language Prediction. J. Cognit. Neurosci. 1–12. https://doi.org/10.1162/jocn_a_01275.

Wicha, N.Y., Bates, E.A., Moreno, E.M., Kutas, M., 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. Neurosci. Lett. 346 (3), 165–168. https://doi.org/10.1016/S0304-3940(03)00599-8.

Wlotko, E.W., Federmeier, K.D., 2012. So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. Neuroimage 62 (1), 356–366.