

# Supplementary Materials

## Supplementary Methods

### 1. Supplementary details about stimuli

*(a) Development of stimuli and cloze norming studies*

*(b) List composition and counterbalancing*

### 2. Spatial search region used for MEG source localization

## Supplementary Results

### 1. Supplementary behavioral results

### 2. Exploratory analysis of ERP and source-localized MEG data between 200-300ms

### 3. MEG source-localization videos

### 4. Exploratory analysis of a subset of the MEG data

### 5. Exploratory analysis of MEG source-level activity over the right hemisphere

## Supplementary Discussion

### 1. The relationship between the MEG and ERP findings

### 2. The retention of dipole polarity when carrying out MEG source localization

*(a) Different conditions can produce dipoles of opposite polarities within the time-windows and anatomical regions*

*(b) The same condition can produce dipoles within the same region that reverse in polarity across time-windows*

*(c) Methodological and theoretical implications*

## Supplementary cross-ROI multivariate decoding in MEG source space

### 1. Motivation

### 2. Methods

### 3. Results and Discussion

## Supplementary Methods

### 1. Supplementary details about stimuli

#### *(a) Development of stimuli and cloze norming studies*

To develop the discourse scenarios, we carried out two cloze norming studies. For both norming studies, participants were recruited through Amazon Mechanical Turk. They were asked to complete each context with the first word that came to mind (Taylor 1953), and in an extension of the standard cloze procedure, to then provide two additional words that could complete the sentence (Schwanenflugel and Lacount 1988; Federmeier et al. 2007). Individuals were excluded if (a) their first language was anything other than English, (b) they self-reported any psychiatric or neurological disorders, or (c) they failed to follow instructions (we included “catch” questions that served as attention checks).

The first cloze norming study aimed to characterize a subset of verbs, which we used to construct the final sentences of each three-sentence scenario. Specifically, this norming study served to establish the lexical constraints and animacy constraints of these verbs in minimal contexts. We began with a large set of 617 transitively-biased verbs that were taken from a number of different sources, including a set of linguistically-characterized verbs (Levin 1993) and materials from previous studies carried out in our lab (Paczynski and Kuperberg 2011, 2012). We excluded verbs with a log Hyperspace Analogue to Language (HAL) frequency (Lund and Burgess 1996) of two standard deviations below the mean (based on English Lexicon Project database, (Balota et al. 2007)). For each verb, we constructed a minimum context, consisting of a proper name, the verb, and a determiner (e.g., “*Harry explored the...*”). For cloze norming, we divided these sentence stems into six lists in order to reduce time demands on any individual. After exclusions, between 89 and 106 participants provided completions for each item.

Based on the animacy of the noun completions, we categorized the verbs as either animate constraining or inanimate constraining, and tallied the number of participants who produced the best completions in order to calculate the lexical constraints of each verb in minimal contexts. We then selected a subset of these verbs (50% animate constraining; 50% inanimate constraining), with lexical constraints less than 24%.

For each verb, we then wrote a corresponding discourse context. Each discourse context consisted of two introductory sentences, and a third sentence that included an adjunct phrase (1-4 words), a pronominal subject that referred back to the first two sentences, the verb and a determiner. We quantified the constraint of these discourse contexts by carrying out a second cloze norming study. In this study, lists were divided into thirds to minimize time demands on any individual participant. After exclusions, between 51 and 69 participants provided completions for each context.

*(b) List composition and counterbalancing*

In the main manuscript, we report contrasts between three experimental conditions: *Expected*, *Unexpected plausible* and *Implausible* in which the critical words followed high constraint contexts. We created 100 *expected* discourse scenarios by pairing 100 high constraint discourse contexts with the noun that was produced most frequently in the second cloze norming study. We created 100 *unexpected plausible* scenarios by pairing the same high constraint contexts with direct object nouns of low cloze values but that were still plausible in context. Finally, we created 150 *implausible* scenarios in which the noun violated the animacy constraints of the prior verb. In 100 of these *implausible* scenarios, we used the same high constraint contexts as the other two conditions. The other 50 *implausible* scenarios were created using 50 additional high constraint discourse contexts. These had served as fillers in our previous ERP study (Kuperberg et al. 2020), but, in the current study, we made the *a priori* decision to include them in our analyses so that we could maximize our statistical power to detect underlying neuroanatomical sources.

In addition to these 350 experimental scenarios, we also constructed an additional 350 scenarios with low constraint contexts. Of these additional low constraint scenarios, 150 were plausible and constituted a fourth *low constraint unexpected* condition (see Kuperberg et al. 2020). Although we do not report analyses involving these stimuli in the present manuscript, they were considered a fourth condition for the purpose of counterbalancing, and we plan to report analyses that include this condition in a separate manuscript. The remaining 200 low constraint scenarios were implausible and were considered fillers. These fillers were based on the *low constraint anomalous* condition that was included in our previous ERP study (Kuperberg et al. 2020).

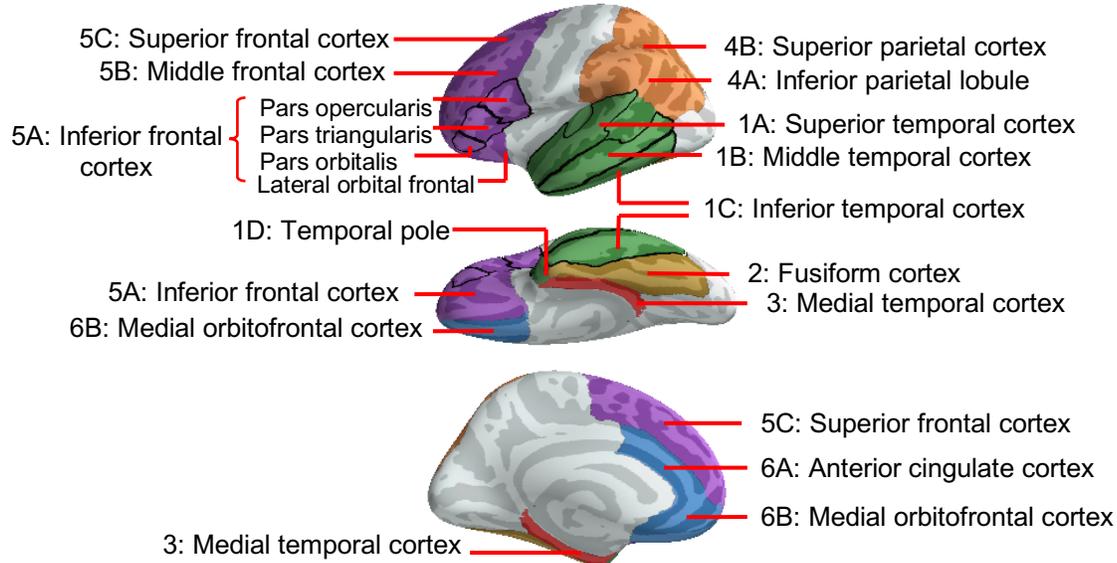
The full set of scenarios were divided into four lists, which were rotated across participants. Each list included 200 scenarios in total: our 100 experimental scenarios in which critical words

followed high constraint contexts (25 *Expected*, 25 *Unexpected plausible*, and 50 *Implausible*), and 100 additional scenarios in which critical words followed low constraint contexts (50 *Low constraint plausible* and 50 *Low constraint implausible*). Therefore, each list contained the same number of plausible and implausible critical nouns, and the same number of high constraint and low constraint discourse contexts. In creating the lists, we were able to partially counterbalance the introductory two sentences of each scenario, the verbs that preceded the critical nouns in the third sentence, and critical nouns (across the *unexpected plausible* and *implausible* conditions) so that participants saw these aspects of the scenario only once, but across all participants, they were seen in more than one condition.

## 2. Spatial search region used for MEG source localization

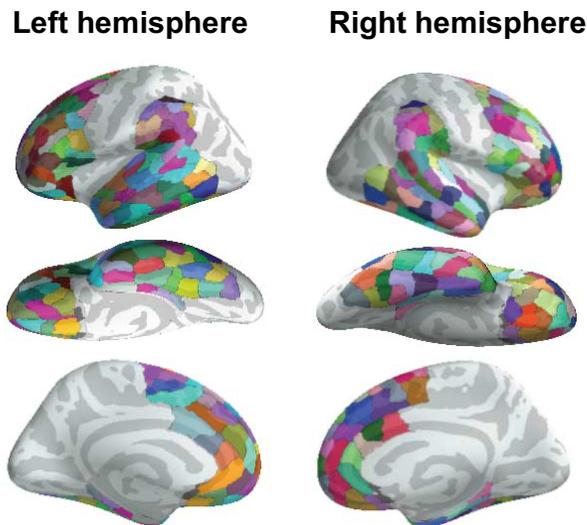
### Supplementary Figure 1: Search region

#### A.



**A. Left-lateralized search region used to carry out MEG statistical analysis.** These regions were defined on the “fsaverage” FreeSurfer surface (Fischl et al. 1999) using the Desikan–Killiany atlas (Desikan et al. 2006). Supplementary Table 1 lists the correspondence between the names of the regions indicated here, and the nomenclature of the equivalent regions in the Desikan–Killiany atlas.

#### B.



**B. Patches on the left and right “fsaverage” FreeSurfer surfaces used for spatial cluster-based permutation tests.** Over the left hemisphere, 140 patches were defined within our *a priori* search area, and were used to account for multiple spatial comparisons (results reported in main manuscript). Over the right hemisphere, 141 patches were defined, and were used to account for multiple spatial comparisons in an exploratory analysis, reported here in Supplementary Materials.

**Supplementary Table 1. A detailed description of the left lateralized search region used for MEG analysis.**

Region	Number	Desikan-Killiany atlas
<i>Lateral temporal cortex</i>		
#Superior temporal cortex	1A	superiortemporal-lh bankssts-lh transversetemporal-lh
Middle temporal cortex	1B	middletemporal-lh
Inferior temporal cortex	1C	inferiortemporal-lh
Temporal pole	1D	temporalpole-lh
<i>Ventral temporal cortex</i>		
Fusiform cortex	2	fusiform-lh
<i>Medial temporal cortex</i>		
Medial temporal cortex	3	parahippocampal-lh entorhinal-lh
<i>Lateral parietal cortex</i>		
Inferior parietal lobule	4A	inferiorparietal-lh supramarginal-lh
Superior parietal cortex	4B	superiorparietal-lh
<i>Lateral frontal cortex</i>		
^Inferior frontal cortex	5A	parsorbitalis-lh parstriangularis-lh parsopercularis-lh lateralorbitofrontal-lh frontalpole-lh
Middle frontal cortex	5B	caudalmiddlefrontal-lh rostralmiddlefrontal-lh
*Superior frontal cortex	5C	superiorfrontal-lh
<i>Medial frontal cortex</i>		
** Anterior cingulate cortex	6A	rostralanteriorcingulate-lh caudalanteriorcingulate-lh
Medial orbitofrontal cortex	6B	medialorbitofrontal-lh

The names and numbers of each region correspond to those indicated in Supplementary Figure 1, which illustrates the full search region. Regions were defined on the “fsaverage” FreeSurfer surface (Fischl *et al.* 1999), using the Desikan–Killiany atlas (Desikan *et al.* 2006). This table lists the correspondences between the numbers and names of the regions shown in Supplementary Figure 1 and the names of the regions from the Desikan–Killiany atlas.

#We grouped multiple gyri defined in the Desikan-Killiany atlas into one cortical region. The gyrus includes the part visible on the pial view plus its adjacent banks of the sulci delineating this gyrus.

^The lateral portion of the orbitofrontal cortex and frontal pole are included in the left inferior frontal cortex.

\*Both lateral and medial surfaces are included within the superior frontal region.

\*\*Both anterior and middle surfaces are included within the anterior cingulate cortex.

## Supplementary Results

### 1. Supplementary behavioral results

On average, participants provided accurate plausibility judgments on 88.3% of trials (SD: 11.1%). There was a significant main effect of Scenario Type,  $F(1,31) = 45.85, p < 0.001, \eta^2 = 0.60$ . Follow-up pairwise comparisons indicated that participants were most accurate in responding “Yes” to the *expected* scenarios (Mean: 95.6%; SD: 5.6%), followed by responding “No” to the *implausible* scenarios (Mean: 88.3%, SD: 8.0%), and they were least accurate in responding “Yes” to the *unexpected plausible* scenarios (Mean: 81.0%, SD: 13.0%). On average, 80.4% of the comprehension questions were answered correctly (SD: 14.2%), which suggests that participants were attending to both the introductory context and the final critical sentence.

### 2. Exploratory analysis of ERP and source-localized MEG data between 200-300ms

As shown in Figure 3A in the main manuscript, there also appeared to be a divergence between the ERP waveforms evoked by the *expected* words and the two other conditions between 200-300ms (before the 300-500ms N400 time-window). This effect had a broad scalp distribution that was maximal over anterior and central electrodes. To explore this effect further, we carried out a post-hoc analysis by collapsing across all spatial regions and all time-points between 200-300ms, and carrying out pairwise statistical between our three conditions of interest. This analysis confirmed a significant difference between the *expected* and both the *unexpected plausible* ( $t_{(31)} = 3.62, p < .001, d = 1.30$ ) and the *implausible* ( $t_{(31)} = 5.75, p < .001, d = 2.07$ ) words, but no difference between the *unexpected plausible* and the *implausible* words ( $t_{(31)} = 0.99, p = 0.33, d = 0.36$ ).

This finding replicates previous ERP studies that have also observed effects of predictability *before* the classical 300-500ms (N400) time-window. The precise interpretation of this and other early effects of predictability during language comprehension remain unclear (see Nieuwland 2019 for a review). One possibility is that the effect reflected an *enhanced positivity* to *expected* words that *confirmed* prior predictions due to an early P3b that reflected the successful categorization of a confirmed prediction (Roehm et al. 2007; Molinaro and Carreiras 2010; Vespignani et al. 2010), or an enhanced P2 due to a top-down attentionally-mediated extraction of visual features (Federmeier et al. 2005). This attention explanation would be consistent with the

theory that activity to expected inputs is initially enhanced before it is suppressed (see Press et al. 2020; see also Feldman and Friston 2010).

Another possibility that has been discussed in the prior psycholinguistic literature is that this early 200-300ms ERP effect reflected an early *reduced negativity* to the expected incoming words – an effect that was superimposed upon the overall positive-going waveform in this time window. This reduced negativity to expected words might have reflected (a) an early divergence of the N400 component itself (see Lau, Holcomb, et al. 2013 for discussion), (b) a reduced N250 (Lau, Holcomb, *et al.* 2013; Brothers et al. 2015), which is thought to reflect sublexical orthographic processing (Holcomb and Grainger 2006; Kiyonaga et al. 2007; Grainger and Holcomb 2009), or (c) a reduced phonological mismatch negativity (PMN), reflecting facilitated phonemic processing (Connolly and Phillips 1994; see also van den Brink et al. 2001). Within a predictive coding framework, an early facilitation effect to *expected* words would correspond to the reduction of sublexical (e.g., orthographic/phonological) prediction error. On this account, in high constraint contexts, the brain not only generates top-down lexico-semantic predictions that suppress prediction error produced by expected inputs at the lexico-semantic level between 300-500ms, but it also generates predictions at sublexical levels of representation that suppress prediction error produced by expected inputs at these lower levels of linguistic representation between 200-300ms.

Given spatiotemporal overlap amongst multiple positive- and negative-going ERP components in this early 200-300ms time-window, our ERP results alone cannot distinguish between these possibilities. However, as shown in Figure 5 (left), the MEG source-localized dynamic Statistical Parametric Maps (dSPMs) within this 200-300ms time-window revealed effects that were more compatible with a reduction (rather than an enhancement) of neural activity to the *expected* words. Specifically, the *expected* words appear to show *less* activity than both the *unexpected plausible* and the *implausible* words in superior temporal, medial temporal and posterior (occipitotemporal) fusiform cortices.

On the other hand, these MEG findings do not shed light on whether these reductions in activity reflect facilitation at the lexico-semantic or sublexical levels of representation. Moreover, paired statistical contrasts revealed significant effects only within superior and medial temporal cortices, and only when comparing the *implausible* and *expected* inputs. The failure to find statistical effects in other regions and/or other contrasts may be due to a lack of power: the ERP

effect between 200-300ms is much smaller than the later N400 effect. Therefore, in order to accurately source-localize this early effect, it will be important to carry out an MEG study that has a very large number of items per condition, and to fully counterbalance the same critical words across levels of predictability. Such a study would provide important data for resolving controversies regarding the validity and functional interpretation of early predictability effects.

### 3. MEG source-localization videos

In the attached videos, we show averaged dSPM source activations from 0-1000ms, in 10ms bins, after critical word onset in each of the three experimental conditions.

### 4. Exploratory analysis of a subset of the MEG data

As noted in the main manuscript, the critical words in the *expected* scenarios were more frequent and had smaller orthographic neighborhoods than the critical words in the two other conditions. These lexical differences between the *expected* words and the other conditions were a function of our counterbalancing scheme, which required us to use the same high constraint discourse contexts across conditions. In addition, the *implausible* condition included twice as many trials as the two other conditions (50 *versus* 25). This ensured that each participant saw an equal proportion of plausible and implausible scenarios. We made the *a priori* decision to include all 50 implausible scenarios in our main analyses in order to maximize power.

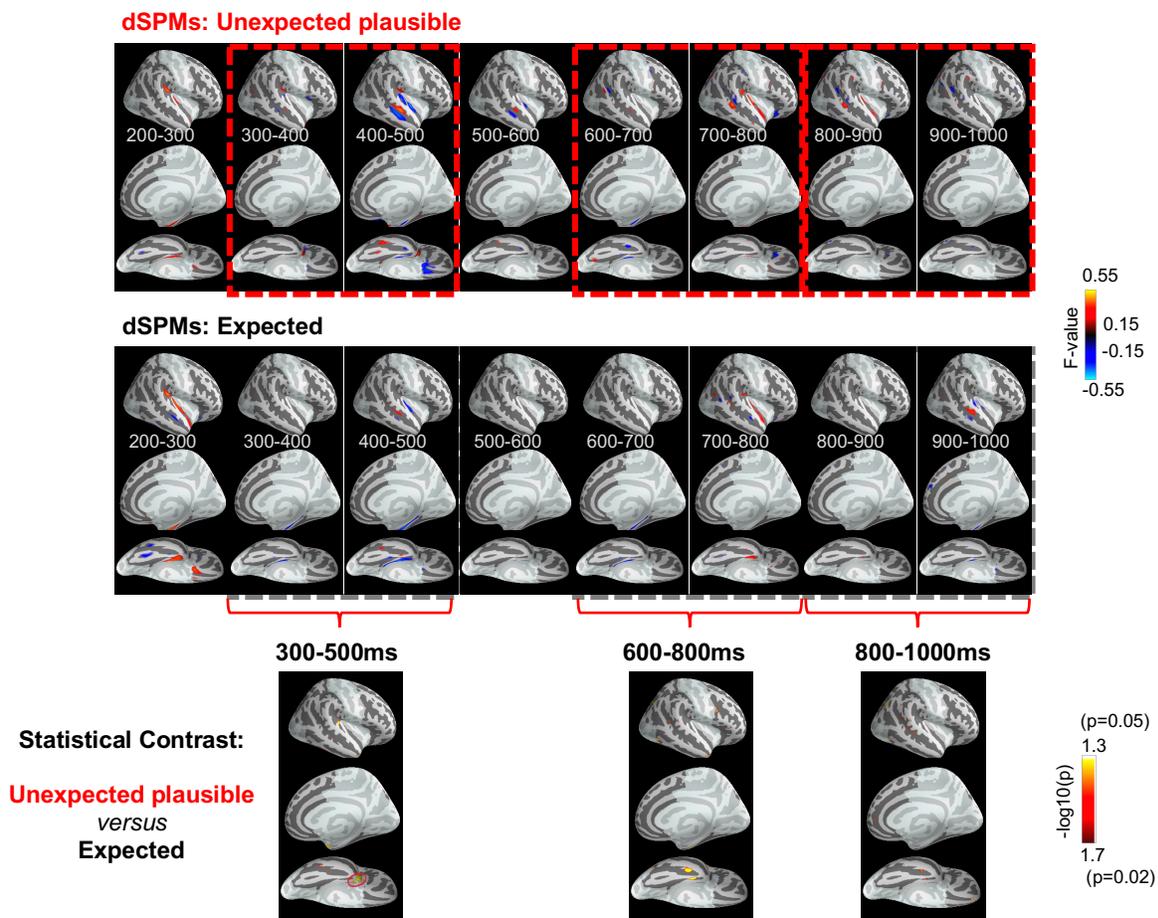
We think that it is unlikely that either of these factors systematically influenced our results. First, although previous ERP studies have shown that both frequency and orthographic neighborhood can modulate the N400 (Rugg 1990; Young and Rugg 1992; Holcomb et al. 2002; Barber et al. 2004; Laszlo and Federmeier 2011, 2014), these lexical effects tend to be much smaller than the effects of predictability and contextual plausibility on the N400. Second, we used *non-parametric* mass univariate statistical tests to analyze our data, which are robust to differences in numbers of trials between conditions (the null distribution was created by randomly permutating the same dataset with the same signal-to-noise ratio).

Nonetheless, to alleviate any concerns that either unmatched lexical properties or unequal trial numbers influenced our results, we conducted an additional analysis. In this analysis, within each participant, we randomly selected a subset of 25 *implausible* scenarios with critical words that matched the *expected* critical words on lexical characteristics (i.e. word length, word

frequency and orthographic neighborhood size). We then computed the ERFs of these randomly selected *implausible* trials within each participant, and compared the source-localized activity within the 300-500ms time-window with activity produced by the *expected* trials across all participants. This analysis revealed the same pattern of results as that reported in the main manuscript.

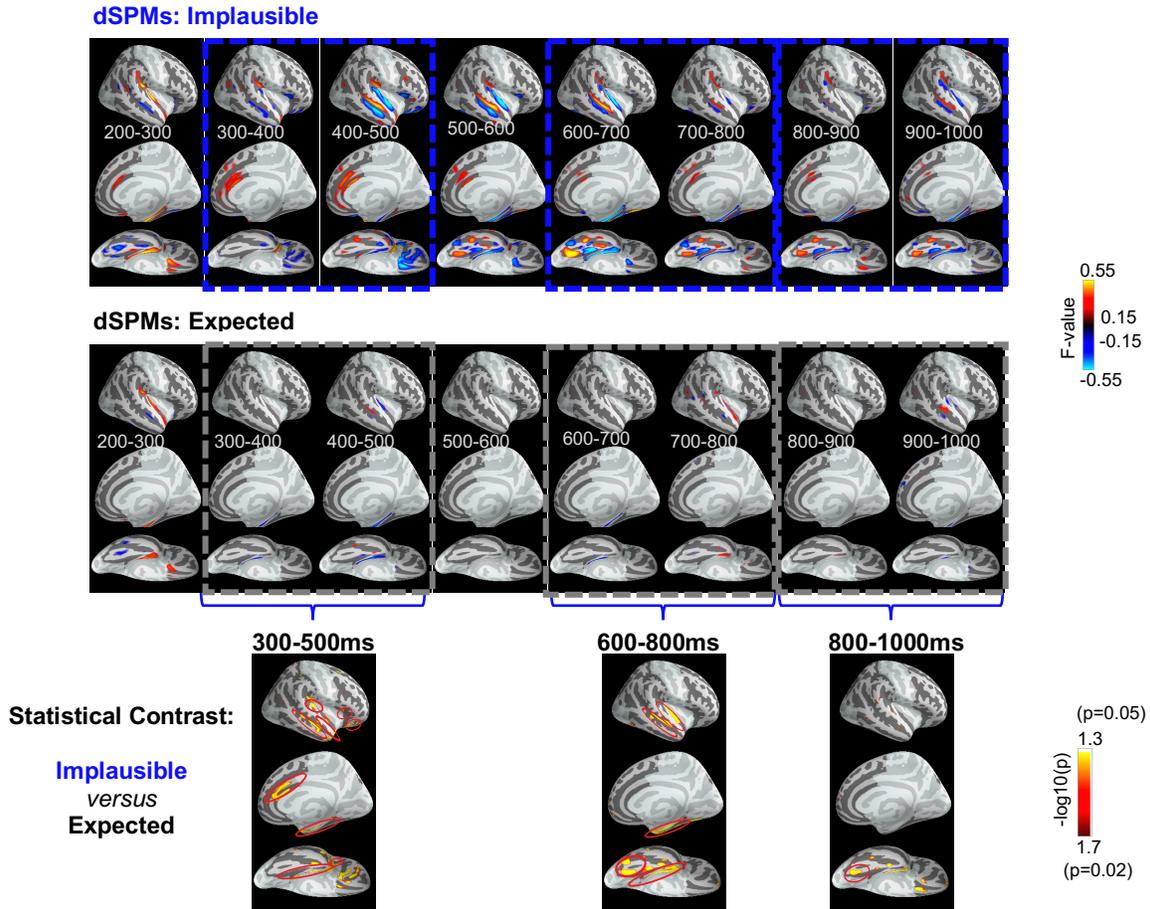
### 5. Exploratory analysis of MEG source-level activity over the right hemisphere

#### Supplementary Figure 2. Exploratory analysis of MEG source-level activity over the right hemisphere produced by the unexpected plausible and the expected critical words.



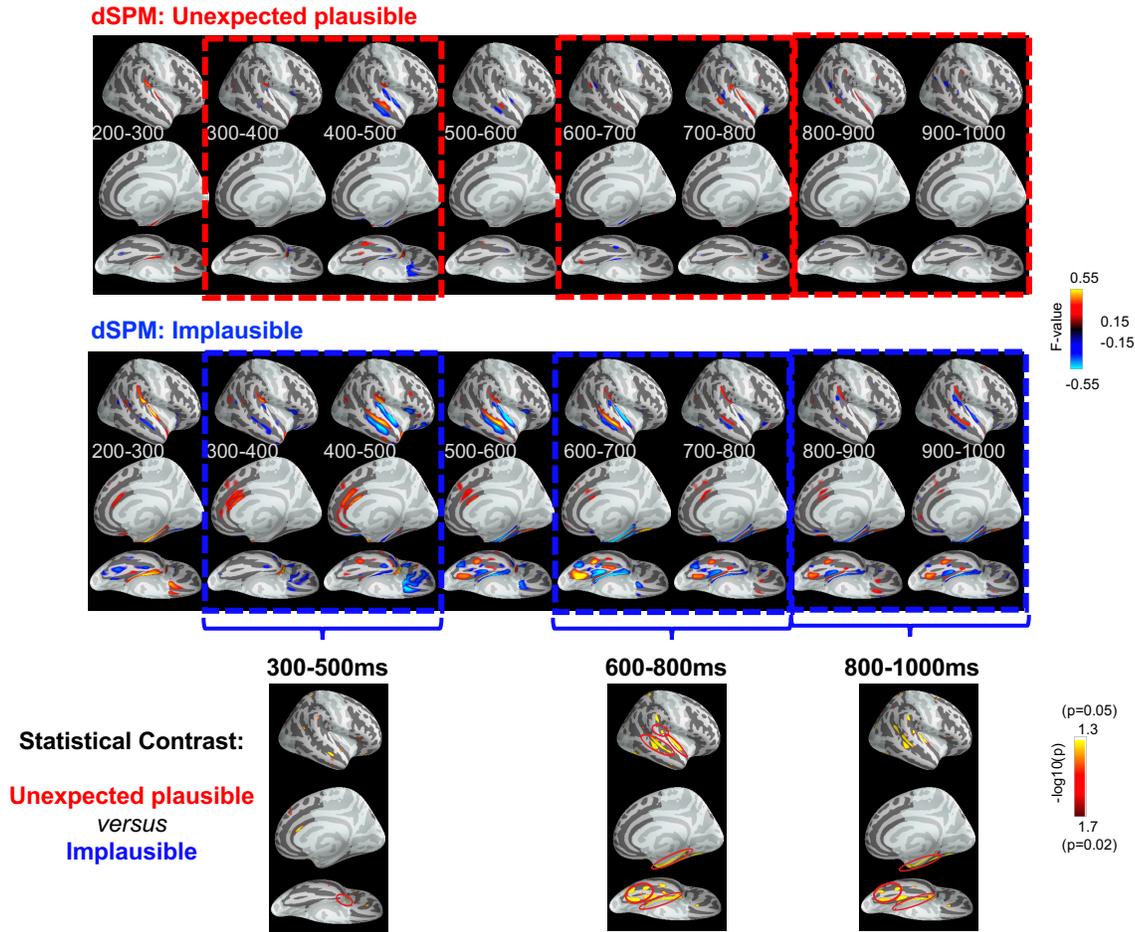
*Top and middle:* Signed dynamic Statistical Parametric Maps (dSPMs) produced by the *unexpected plausible* and the *expected* critical words, shown at 100ms intervals from 200ms until 1000ms. All dSPMs are thresholded at 0.15, with red indicating outgoing currents and blue indicating ingoing currents. *Bottom:* Statistical maps contrasting the *unexpected plausible* and *expected* critical words within our three time-windows of interest: 300-500ms, 600-800ms, and 800-1000ms. Red circles indicate regions that reached cluster-level significance. Within the 300-500ms time window, the *unexpected plausible* critical words evoked significantly more activity than the *expected* critical words within the right temporal pole. No significant effects were found between 600-1000ms.

**Supplementary Figure 3. Exploratory analysis of MEG source-level activity over the right hemisphere produced by the implausible and the expected critical words.**



*Top and middle:* Signed dynamic Statistical Parametric Maps (dSPMs) produced by the *implausible* and the *expected* critical words, shown at 100ms intervals from 200ms until 1000ms. All dSPMs are thresholded at 0.15, with red indicating outgoing currents and blue indicating ingoing currents. *Bottom:* Statistical maps contrasting the *implausible* and *expected* critical words within our three time-windows of interest: 300-500ms, 600-800ms, and 800-1000ms. Red circles indicate regions that reached cluster-level significance. Within the 300-500ms time window, the *implausible* critical words evoked significantly more activity than the *expected* critical words within the right lateral temporal cortex, the right anterior inferior frontal cortex, and the right anterior cingulate cortex. This contrast also revealed an effect in the right medial temporal cortex, which was driven by dipoles going in opposite directions to the *implausible* (outgoing) and the *expected* (ingoing) critical words. The locations of these effects were similar to those observed over the left hemisphere, but they appeared to be less robust. Between 600-1000ms, the *implausible* critical words produced more activity within the right posterior fusiform cortex than the *expected* critical words (significant in both the 600-800ms and the 800-1000ms windows). Between 600-800ms, the *implausible* critical words also produced significantly more activity than the *expected* critical words within the right superior temporal and medial temporal cortices.

**Supplementary Figure 4. Exploratory analysis of MEG source-level activity over the right hemisphere produced by the unexpected plausible and the implausible critical words.**



*Top and middle:* Signed dynamic Statistical Parametric Maps (dSPMs) produced by the *unexpected plausible* and *implausible* critical words, shown at 100ms intervals from 200ms until 1000ms. All dSPMs are thresholded at 0.15, with red indicating outgoing currents and blue indicating ingoing currents. *Bottom:* Statistical maps contrasting the *unexpected plausible* and *implausible* critical words within our three *a priori* time windows of interest: 300-500ms, 600-800ms, and 800-1000ms. Red circles indicate regions that reached cluster-level significance. Within the 300-500ms time window, the *implausible* critical words evoked significantly more activity than the *unexpected plausible* critical words within the right temporal pole. Between 600-1000ms, the *implausible* critical words produced more activity within the right posterior fusiform and medial temporal regions than the *unexpected plausible* critical words (significant in both the 600-800ms and the 800-1000ms windows). Between 600-800ms, *implausible* critical words also produced significantly more activity within the right superior temporal cortex than the *unexpected plausible* critical words.

## Supplementary Discussion

### 1. The relationship between the MEG and ERP findings

As noted in the main manuscript, although MEG and ERP both measure phase-locked evoked activity, they do not capture precisely the same underlying signal, and so they do not always pattern together (Ahlfors, Han, Belliveau, et al. 2010). By simultaneously collecting ERP and MEG data using the same stimuli in the same participants, we were able to directly compare the ERP and MEG effects. In general, the two methods revealed similar patterns of modulation across the three conditions (e.g. graded increases of activity within the N400 time-window, and spatially distinct effects to the *unexpected plausible* and *implausible* continuations in the late time-window). However, there were also some interesting differences.

First, the effect of contextual plausibility (*implausible* versus *unexpected plausible*) on the ERP N400 (shown in Figure 4, main manuscript) appeared to be much smaller than on the sensor-level MEG N400 (shown in Figure 5, main manuscript). We suggest that this is because the N400 ERP component evoked by the *implausible* critical words was artificially reduced at the scalp surface as a result of spatiotemporal overlap with the subsequent *late posterior positivity/P600* ERP component that was produced by these continuations. The N400 and *late posterior positivity/P600* ERP components both have posterior scalp distributions. However, because they have opposite polarities, they can cancel each other out at the scalp surface (Kuperberg et al. 2007; Brouwer and Crocker 2017). This type of “component overlap” is less of an issue for MEG for two reasons. First, the signal detected by gradiometer MEG sensors does not carry information about the polarity of the underlying dipoles; that is, sensor-level evoked MEG responses reflect the overall magnitude of activity, regardless of the direction of the underlying currents. Therefore, unlike ERP responses, there is no cancellation of the MEG signal at the scalp surface. Second, MEG has a better spatial resolution than EEG because magnetic fields are less distorted than electric fields by the conductivities of the skull and scalp. Therefore, evoked MEG responses that originate from spatially distinct underlying sources are less likely than ERP responses to overlap spatially at the scalp surface within the same time-window.

Note that this account of ERP component overlap implies that the *late posterior positivity/P600* ERP produced by the *implausible* words began within the N400 time window. This, in turn, implies that the conflict that triggered the *late posterior positivity/P600* effect was

evident in the 300-500ms (N400) time window. This early conflict may have been indexed by the anterior cingulate response to the *implausible* (versus *expected*) words, detected by MEG between 300-500ms.

A second difference between the MEG and ERP findings was in the magnitude of the late effects observed between 600-1000ms. In ERPs, the magnitude of these late effects (both the *late frontal positivity* evoked by the *unexpected plausible* continuations and the *late posterior positivity/P600* evoked by the highly *implausible* continuations) were generally larger than the MEG effects for the same contrasts observed within the same late time window. The relative insensitivity of MEG to neural effects that manifest in the ERP waveform as robust positive-going components has been noted before (Ahlfors, Han, Lin, et al. 2010). For example, MEG is relatively insensitive to the well-known domain-general P3b effect (Siedenberg et al. 1996), to which the late posterior positivity/P600 is thought to be functionally related (Coulson et al. 1998; Osterhout et al. 2012; Sassenhagen et al. 2014; Sassenhagen and Fiebach 2019). One possible reason for this is that, unlike ERPs, which index activity originating from both sulci and gyri, MEG is insensitive to radial sources from gyri (Ahlfors, Han, Belliveau, et al. 2010). In addition, in MEG, tangential sources on opposing sides of sulci often cancel out (Ahlfors, Han, Lin, et al. 2010). For both these reasons, MEG is relatively insensitive to activity that stems from *extended* regions of cortex that cut across multiple gyri and sulci, and that may produce large late positivity ERP effects.

These differences between ERP and MEG may also have functional implications; that is, the two measures may not necessarily index precisely the same *cognitive* mechanisms. For example, consider the late EEG and MEG evoked responses produced by the highly *implausible* inputs between 600-1000ms: In EEG, this late response manifests as a posteriorly distributed positive-going ERP component – the well-known semantic P600 (Kolk et al. 2003; Kuperberg et al. 2003; Hoeks et al. 2004; Kim and Osterhout 2005; Kuperberg 2007; Kuperberg et al. 2020). In MEG, the late evoked effect to the semantic errors manifests as increased activity within the fusiform, inferior frontal and medial frontal cortices. This, however, does not necessarily imply that the MEG and EEG evoked responses indexed precisely the same neurocognitive mechanisms. Indeed, we have recently proposed an “informing seeking” theory, inspired by more general computational principles of active information sampling (MacKay 1992; Chater et al. 1998; Nelson 2005; Gottlieb 2012; Gottlieb and Oudeyer 2018), in which prolonged processing in response to

linguistic errors is posited to involve a number of distinct processes that proceed in parallel, in the same time window, interacting closely with one another (Kuperberg et al. In preparation).

According to this information seeking theory, lower-level perceptual reprocessing of a conflicting input serves to reduce the comprehender's uncertainty about whether or not the input was correctly processed the first time around, or, equivalently, to increase their confidence about whether the input actually contained an anomaly, i.e. was a production error (see also van de Meerendonk et al. 2009). This reprocessed input continually feeds into a distinct evidence accumulation mechanism that tracks the brain's uncertainty/confidence in this decision. The P600 itself is proposed to reflect this second-order decision variable that tracks the accumulated evidence for the source of the error at each point in time (functionally analogous to other domain-general late (long-latency) positive-going members of the P300 family that may track evidence about whether a preceding choice was correct (Steinhauser and Yeung 2010; Boldt and Yeung 2015; Murphy et al. 2015; Desender et al. 2019), and see Desender et al. 2021 for a review).

It is therefore possible that the MEG late evoked response to the highly implausible inputs within the posterior fusiform cortex was relatively more sensitive to the more localized process of orthographically reprocessing the input (here, interpreted as a failure to suppress lower-level prediction error), while EEG recordings was more sensitive to decision-making and evidence accumulation mechanisms, reflected by the P600.

## 2. The retention of dipole polarity when carrying out MEG source localization analyses

When carrying out distributed source localization of our MEG data, we chose to retain the polarity (signed values) of the estimated dipoles. This allowed us not only to estimate the magnitude of activity across experimental conditions, but also to determine the direction of the underlying dipoles, i.e. whether the current was ingoing or outgoing, relative to the cortical surface. At a neurophysiological level, the polarity of a dipole is determined both by the precise configuration of the pyramidal cells and the direction of the intracellular currents they generate (Lopes da Silva 2010). Although the precise mechanisms that give rise to differences in dipole polarity are unclear, it is likely that systematic *differences* in dipole polarity between conditions and/or time-windows has some functional significance. As we discussed next, our results revealed some interesting patterns.

*(a) Different conditions can produce dipoles of opposite polarities within the same time-windows and neuroanatomical regions*

In several cases, we found that the statistical differences between two conditions within a given region and time-window were driven by dipoles going in opposite directions to each condition. For example, in the N400 time-window, the effect in the left medial temporal cortex, was driven not only by a dipole to the unpredictable words (both to the *unexpected plausible* and *implausible* words), but also by a dipole to the *expected* words. As discussed in the main manuscript, we speculated that the dipole produced by the *expected* words indexed the detection of a match between pre-activated activity within this medial temporal region, and the expected bottom-up input.

Beyond its theoretical implications, our finding that effects can be driven by dipoles going in opposite directions also has methodological implications. Previous intracranial studies have also reported effects within medial temporal cortex within the N400 time-window, and it has been noted that different electrodes within these medial temporal regions produce local field potentials of opposite polarities to incongruous and expected words (McCarthy et al. 1995). However, previous MEG studies using distributed source localization have failed to report effects within the medial temporal cortex within the N400 time window, either in semantic priming paradigms (Lau, Gramfort, et al. 2013; Lau et al. 2016) or during sentence comprehension (Maess et al. 2006). We suggest that this may be because these previous MEG studies used unsigned rather than signed dipole values for source localization. The absolute values of two dipoles going in opposite directions would have therefore canceled out, and failed to show any significant difference at all to predictable versus unpredictable words.

In addition to the medial temporal effects between 300-500ms, we also observed effects that were driven by dipoles going in opposite directions across conditions within the later 600-1000ms time-window: (1) the effects within left inferior frontal and middle temporal cortices in contrasting the *unexpected plausible* and *expected* words, and (2) the effect within left inferior frontal cortex in contrasting the *implausible* and *expected* words. Once again, these differences would have been overlooked if we had only examined the absolute amplitude of the source activity.

*(b) The same condition can produce dipoles within the same region that reverse in polarity across time-windows*

In several other cases, a dipole produced by a particular condition showed a polarity reversal between the 300-500ms and the later 600-1000ms time-window. This was the case for the evoked responses produced by the *unexpected plausible* words within the left lateral temporal cortex, as well as the response produced by the *implausible* words within the left inferior frontal and medial temporal cortices. Again, although the precise mechanisms underlying these dipole reversals are unclear, they are again likely to have some functional significance. In the main manuscript, we noted that these dipole reversals were consistent with a functional distinction between early feedforward activity within the 300-500ms time-window and later feedback activity within the 600-1000ms time-window.

*(c) Methodological and theoretical implications*

Taken together, these observations add to a growing body of work suggesting that there are several advantages of retaining dipole polarity when estimating underlying neural sources. Previous MEG work that has systematically compared analyses of the same datasets using signed and unsigned estimates, has shown that retaining signed summary statistics is robust to factors such as individual neuroanatomical variability and spatial smoothing (Henson et al. 2007). In addition, a recent study showed that retaining dipole direction information allowed for a better characterization of the time course of visual word processing (Gwilliams et al. 2016). As discussed in the Methods section of the main manuscript, the retention of dipole information also has the advantage of allowing analyses between conditions with unequal numbers of trials, in contrast to traditional methods that square positive and negative values to yield positively-signed estimate of dipole magnitude, thereby artificially inflating noise estimates in conditions with more trials. Finally, as discussed above, retaining dipole polarity ensures that effects that are driven by dipoles going in opposite directions are not canceled out.

The present set of findings suggest that, beyond these methodological implications, a systematic examination of dipole polarity in MEG source localization may also have functional implications, yielding new insights into the cognitive mechanisms that support language comprehension.

## Supplementary cross-ROI multivariate decoding in MEG source space

### 1. Motivation

The primary aim of our study was to neuroanatomically localize the event-related (or evoked) neural activity produced by expected, unexpected plausible and implausible inputs between 300-500ms (the time-window of the N400) and between 600-100ms (the time window corresponding to two late positive-going ERP components). We therefore carried out univariate analyses of both our EEG and MEG data to detect differences in the *magnitude* of neural activity across conditions.

Importantly, however, differences in information content within a cortical region are not always accompanied by differences in the total magnitude of neural activity. There is therefore increasing interest in using *multivariate* methods to decode stimulus- or condition-specific information, which is thought to be encoded as *patterns* of neural activity. In many situations, univariate and multivariate responses are highly correlated; that is, larger univariate responses are often associated with a larger multivariate response (more decodable information). However, univariate and multivariate activity can sometimes dissociate. For example, two regions can both produce increased evoked neural responses, without sharing patterns of decodable multivariate activity. Conversely, two regions can share decodable information, without showing increases in evoked neural activity.

The predictive coding framework described in the main manuscript provides a principled account of why univariate and multivariate activity can sometimes dissociate. As discussed in the main manuscript, larger evoked responses are thought to be primarily produced by increased activity within error units, while decodable information is carried by both error and state units. Here we report a cross-ROI multivariate decoding analysis that aimed to explore whether or not information was shared between left frontal and temporal regions as unpredicted (*vs.* predicted) information was passed up and down the cortical hierarchy. The continuous transfer of information between temporal and inferior frontal cortices has been well-established in previous studies (Mamashli et al. 2019; Liu et al. 2020). Recently, multivariate methods have been used to demonstrate that, between 300-500ms, *both* predicted and unpredicted information is continually shared between temporal and inferior frontal cortices (Lyu et al. 2019). These previous findings

are consistent with most frameworks of language processing (e.g. Baggio and Hagoort 2011), including predictive coding.

In this current exploratory multivariate analysis, we asked whether condition-specific patterns of neural activity in one brain region — patterns that can *discriminate* between unexpected and expected words — resemble condition-specific patterns of activity in another region. In the 300-500ms (N400) time window, if lexico-semantic prediction error is passed up from the left lateral temporal cortex to update state units within the left inferior frontal cortex, we would expect to see shared condition-specific information across these two regions. Importantly, this condition-specific activity should be decodable in the inferior frontal cortex, regardless of whether we observe an increased evoked response within this region. Similarly, between 600-1000ms, if newly generated schema-relevant predictions are passed down from the left inferior frontal cortex to the left lateral temporal cortex, then we should also see shared condition-specific information in this late time-window. In contrast, we do *not* expect to see shared information between left inferior frontal and lateral temporal or posterior fusiform cortices during the processing of highly implausible/anomalous (*vs.* expected) words. When processing implausible continuations in this late time-window, these cortical regions should carry incompatible contextual and lexical information, resulting in minimal cross-region decoding (see Figure 3 in the main text).

This cross-ROI multivariate decoding analysis was therefore intended to provide preliminary multivariate data to clarify or support certain points in our interpretation of the evoked effects, as discussed in the main manuscript.

## 2. Methods

In the 300-500ms time-window, we trained classifiers within a left temporal region of interest (ROI) to discriminate between (a) the *unexpected plausible* and the *expected* continuations, and (b) the *implausible* and *expected* continuations. We then asked whether these classifiers could be used to decode activity produced within a left inferior frontal ROI, i.e. to determine whether the left inferior frontal region shared information with the left temporal region regarding the predictability of the critical words. These ROIs included all vertices that showed differences in evoked activity between the *implausible* and *expected* continuations within this same 300-500ms time window (at  $p \leq 0.05$  uncorrected).

In the 600-1000ms time-window, we first trained classifiers within a left inferior frontal ROI to discriminate between the *unexpected plausible* and the *expected* continuations. We then asked whether these classifiers could be used to decode activity produced within a left temporal ROI. The left inferior frontal and left temporal ROIs were defined based on the difference in evoked responses between the *unexpected plausible* and *expected* words within this same 600-1000ms time window (at  $p \leq 0.05$  uncorrected). In addition, we trained classifiers within a left inferior frontal ROI to discriminate between the *implausible* and *expected* continuations, and tested whether these classifiers could be used to decode activity produced within a left posterior fusiform ROI. Here, the left inferior frontal and left posterior fusiform ROIs were defined based on the difference in evoked responses between the *implausible* and *expected* words within the 600-1000ms time window.

For each analysis, we extracted single-trial MEG activity from individual vertices within the relevant pairs of spatial ROIs. To do this, we applied each participant's inverse operator to their single-trial sensor-level MEG data (at all magnetometer and gradiometer sensors). This yielded a matrix with dimensions of *trial*  $\times$  *vertex*  $\times$  *time*. To ensure that activity within each pair of ROIs was in the same dimensional space, we rotated these matrices in the direction of maximum variance using Principal Component Analysis (PCA), retaining the top components whose variance summed to at least 99% across all time samples.

Using logistic regression, we first trained classifiers to distinguish activity between the two conditions that comprised a given contrast at each time point within one of the two ROIs, and then asked whether these trained classifiers were able to discriminate activity between the same two conditions at corresponding time points within its paired ROI. We conducted the cross-ROI decoding analysis on these PCA-transformed data using the *sklearn* package in python. We averaged the cross-ROI decoding accuracy values across all time points within the time-window of interest and used a one-sample t-test to determine whether this average decoding accuracy was significantly different from 50% (chance level).

### 3. Results and Discussion

#### 300-500ms

##### *Unexpected plausible vs. expected*

As reported in the main manuscript, this contrast revealed a univariate N400 effect that localized only to the left temporal cortex. The multivariate cross-temporal coding analysis nonetheless showed that, within this same 300-500ms time window, *unexpected* (vs. *expected*) information was shared between the left temporal and left inferior frontal cortex — there was a small but significant cross-temporal decoding effect for this contrast (Mean = 51.8%, SD = 5.2%,  $t_{(31)} = 2.14$ ,  $p = 0.041$ ,  $d = 0.77$ ).

This dissociation between the univariate and multivariate findings can be explained within a predictive coding framework. Similar to other frameworks, predictive coding posits that there is continuous exchange of unpredicted information between left temporal and inferior frontal regions (encoded within state units) within the 300-500ms time window (see Baggio and Hagoort 2011, and see Lyu *et al.* 2019; Mamashli *et al.* 2019; Liu *et al.* 2020 for recent empirical evidence). However, within the left inferior frontal cortex, any activity produced by the *unexpected plausible* continuations within higher-level error units would have been suppressed by predictions based on longer-term real-world knowledge, explaining the absence of an enhanced evoked response within this region.

We note, however, that the decoding effect for this contrast was small. This is not surprising given our limited power: We were only looking at decodable activity that distinguished the unpredicted and expected continuations, unlike previous approaches (e.g. Lyu *et al.* 2019). It will therefore be important to replicate this effect using larger numbers of *unexpected* trials and larger numbers of participants.

### *Implausible vs expected*

As reported in the main manuscript, this contrast revealed an N400 effect that localized both to the left temporal and left inferior frontal cortex. The multivariate cross-temporal coding analysis showed that within this same 300-500ms time window, *implausible* (vs. *expected*) information was shared between these two regions. This resulted in a highly significant cross-temporal decoding effect (Mean = 52.9%, SD = 4.9%,  $t_{(31)} = 3.27$ ,  $p = 0.003$ ,  $d = 1.17$ ).

Again, this result is naturally explained within a predictive coding framework: *unexpected implausible* information was passed up from left temporal (encoded within error units) to inferior frontal regions (encoded within state units) within the 300-500ms time window. However, unlike the *unexpected plausible* continuous, the *implausible* information encoded at the inferior frontal

region was not suppressed by predictions based on longer-term real-world knowledge, resulting in enhanced evoked response within this region. The presence of a large prediction error/evoked response in *both* the temporal and frontal regions (decodable information in *both* error and state units) may explain why the cross-ROI decoding effect for this contrast (*implausible vs. expected*) was more robust than the decoding effect observed for the *unexpected plausible vs. expected* contrast, described above.

### 600-1000ms

#### *Unexpected plausible vs. expected*

As reported in the main manuscript, in this later time window, this contrast revealed evoked effects that localized to both left inferior frontal and temporal cortices (the effect in the temporal cortex was of the opposite polarity to that seen in the N400 time window). The cross-ROI multivariate coding analysis showed that, within this same 600-1000ms time window, the *unexpected (vs. expected)* information was shared between these two regions — there was a small but significant cross-temporal decoding effect for this contrast (Mean = 52.0%, SD = 4.9%,  $t_{(31)} = 2.12$ ,  $p = 0.042$ ,  $d = 0.76$ ).

This is consistent with the claims of predictive coding, which interprets these late evoked responses as top-down error (i.e. newly retrieved information within predictions that is not present within the prior states) that is propagated down the cortical hierarchy within this time window. Again, however, we note that the cross-temporal decoding effect was small, and so it will be important to replicate this finding in future studies.

#### *Implausible vs. expected*

In this later time window, we observed evoked effects that localized to the left inferior frontal and fusiform cortices. The multivariate cross-temporal coding analysis, however, showed *no* evidence that the *implausible (vs. expected)* information was shared across these two regions within this same 600-1000ms time window (Mean = 49.4%, SD = 5.3%,  $t_{(31)} = -0.66$ ,  $p = 0.512$ ,  $d = -0.23$ ). Nor did it show any evidence that information was shared between the left inferior frontal and temporal cortex within this time window (Mean = 50.5%, SD = 6.0%,  $t_{(31)} = 0.44$ ,  $p = 0.661$ ,  $d = 0.16$ ).

Again, this is consistent with our interpretation of the evoked effects within a predictive coding framework. These implausible continuations were anomalous (e.g., “*cautioned the \*drawers*”), and therefore *conflicted* with the state of the hierarchical generative model as a whole, given its parameters. We suggest that the late evoked response within the fusiform cortex reflected a late orthographic prediction error – the failure to converge on a single, stable interpretation at the higher levels of cortex, resulted in a failure to switch off this lower-level prediction error. However, there was no evidence of cross-ROI decoding due to the lack of consistent information across the cortical hierarchy.

## References

- Ahlfors SP, Han J, Belliveau JW, Hämäläinen MS. 2010. Sensitivity of MEG and EEG to source orientation. *Brain Topogr.* 23:227-232.
- Ahlfors SP, Han J, Lin FH, Witzel T, Belliveau JW, Hämäläinen MS, Halgren E. 2010. Cancellation of EEG and MEG signals generated by extended and distributed sources. *Hum Brain Mapp.* 31:140-149.
- Baggio G, Hagoort P. 2011. The balance between memory and unification in semantics: A dynamic account of the N400. *Lang Cogn Process.* 26:1338-1367.
- Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R. 2007. The English lexicon project. *Behav Res Methods.* 39:445-459.
- Barber H, Vergara M, Carreiras M. 2004. Syllable-frequency effects in visual word recognition: evidence from ERPs. *Neuroreport.* 15:545-548.
- Boldt A, Yeung N. 2015. Shared neural markers of decision confidence and error detection. *J Neurosci.* 35:3478-3484.
- Brothers T, Swaab TY, Traxler MJ. 2015. Effects of prediction and contextual support on lexical processing: prediction takes precedence. *Cognition.* 136:135-149.
- Brouwer H, Crocker MW. 2017. On the proper treatment of the N400 and P600 in language comprehension. *Front Psychol.* 8:1327.
- Chater N, Crocker MW, Pickering MJ. 1998. The rational analysis of inquiry: The case of parsing. In: Oaksford M, Chater N, editors. *Rational Models of Cognition* New York: Oxford University Press p 441-468.
- Connolly JF, Phillips NA. 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *J Cogn Neurosci.* 6:256-266.
- Coulson S, King JW, Kutas M. 1998. Expect the unexpected: Event-related brain responses to morphosyntactic violations. *Lang Cogn Process.* 13:21-58.
- Desender K, Murphy P, Boldt A, Verguts T, Yeung N. 2019. A post-decisional neural marker of confidence predicts information-seeking in decision-making. *J Neurosci.* 39:3309-3319.
- Desender K, Ridderinkhof KR, Murphy P. 2021. Understanding neural signals of post-decisional performance monitoring: An integrative review. *Elife.* 10.
- Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage.* 31:968-980.
- Federmeier KD, Mai H, Kutas M. 2005. Both sides get the point: hemispheric sensitivities to sentential constraint. *Mem Cognit.* 33:871-886.
- Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. 2007. Multiple effects of sentential constraint on word processing. *Brain Res.* 1146:75-84.
- Feldman H, Friston KJ. 2010. Attention, uncertainty, and free-energy. *Front Hum Neurosci.* 4:215.
- Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage.* 9:195-207.
- Gottlieb J. 2012. Attention, learning, and the value of information. *Neuron.* 76:281-295.
- Gottlieb J, Oudeyer PY. 2018. Towards a neuroscience of active sampling and curiosity. *Nat Rev Neurosci.* 19:758-770.

- Grainger J, Holcomb PJ. 2009. Watching the word go by: On the time-course of component processes in visual word recognition. *Lang Linguist Compass*. 3:128-156.
- Gwilliams L, Lewis GA, Marantz A. 2016. Functional characterisation of letter-specific responses in time, space and current polarity using magnetoencephalography. *Neuroimage*. 132:320-333.
- Henson RN, Mattout J, Singh KD, Barnes GR, Hillebrand A, Friston K. 2007. Population-level inferences for distributed MEG source localization under multiple constraints: application to face-evoked fields. *Neuroimage*. 38:422-438.
- Hoeks J, Stowe LA, Doedens G. 2004. Seeing words in context: The interaction of lexical and sentence level information during reading. *Brain Res Cogn Brain Res*. 19:59-73.
- Holcomb PJ, Grainger J. 2006. On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*. 18:1631-1643.
- Holcomb PJ, Grainger J, O'Rourke T. 2002. An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*. 14:938-950.
- Kim A, Osterhout L. 2005. The independence of combinatory semantic processing: Evidence from event-related potentials. *J Mem Lang*. 52:205-225.
- Kiyonaga K, Grainger J, Midgley K, Holcomb PJ. 2007. Masked cross-modal repetition priming: An event-related potential investigation. *Lang Cogn Process*. 22:337-376.
- Kolk HHJ, Chwilla DJ, van Herten M, Oor PJ. 2003. Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain Lang*. 85:1-36.
- Kuperberg GR. 2007. Neural mechanisms of language comprehension: Challenges to syntax. *Brain Res*. 1146:23-49.
- Kuperberg GR, Alexander E, Brothers T. In preparation. The P600 reflects reanalysis – not error correction.
- Kuperberg GR, Brothers T, Wlotko E. 2020. A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*. 32:12-35.
- Kuperberg GR, Kreher DA, Sitnikova T, Caplan DN, Holcomb PJ. 2007. The role of animacy and thematic relationships in processing active English sentences: evidence from event-related potentials. *Brain Lang*. 100:223-237.
- Kuperberg GR, Sitnikova T, Caplan D, Holcomb PJ. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Brain Res Cogn Brain Res*. 17:117-129.
- Laszlo S, Federmeier KD. 2011. The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*. 48:176-186.
- Laszlo S, Federmeier KD. 2014. Never seem to find the time: Evaluating the physiological time course of visual word recognition with regression analysis of single-item event-related potentials. *Lang Cogn Neurosci*. 29:642-661.
- Lau EF, Gramfort A, Hämäläinen MS, Kuperberg GR. 2013. Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *J Neurosci*. 33:17174-17181.
- Lau EF, Holcomb PJ, Kuperberg GR. 2013. Dissociating N400 effects of prediction from association in single-word contexts. *J Cogn Neurosci*. 25:484-502.

- Lau EF, Weber K, Gramfort A, Hämäläinen MS, Kuperberg GR. 2016. Spatiotemporal signatures of lexico-semantic prediction. *Cerebral Cortex*. 26:1377-1387.
- Levin B. 1993. *English Verb Classes And Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Liu Z, Shu S, Lu L, Ge J, Gao JH. 2020. Spatiotemporal dynamics of predictive brain mechanisms during speech processing: an MEG study. *Brain Lang*. 203:104755.
- Lopes da Silva FH. 2010. Electrophysiological basis of MEG signals. In: Hansen PC, Kringelbach ML, Salmelin R, editors. *MEG: an Introduction to Methods* New York, NY: Oxford University Press p 1-23.
- Lund K, Burgess C. 1996. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods: Instruments & Computers*. 28:203-208.
- Lyu B, Choi HS, Marslen-Wilson WD, Clarke A, Randall B, Tyler LK. 2019. Neural dynamics of semantic composition. *Proc Natl Acad Sci U S A*. 116:21318-21327.
- MacKay DJC. 1992. Information-based objective functions for active data selection. *Neural Computation*. 4:590-604.
- Maess B, Herrmann CS, Hahne A, Nakamura A, Friederici AD. 2006. Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Res*. 1096:163-172.
- Mamashli F, Khan S, Obleser J, Friederici AD, Maess B. 2019. Oscillatory dynamics of cortical functional connections in semantic prediction. *Hum Brain Mapp*. 40:1856-1866.
- McCarthy G, Nobre AC, Bentin S, Spencer DD. 1995. Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. *J Neurosci*. 15:1080-1089.
- Molinaro N, Carreiras M. 2010. Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biol Psychol*. 83:176-190.
- Murphy PR, Robertson IH, Harty S, O'Connell RG. 2015. Neural evidence accumulation persists after choice to inform metacognitive judgments. *eLife*. 4.
- Nelson JD. 2005. Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*. 112:979-999.
- Nieuwland MS. 2019. Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neurosci Biobehav Rev*. 96:367-400.
- Osterhout L, Kim A, Kuperberg GR. 2012. The neurobiology of sentence comprehension. In: Spivey M, Joannisse M, McRae K, editors. *The Cambridge Handbook of Psycholinguistics* Cambridge: Cambridge University Press p 365-389.
- Paczynski M, Kuperberg GR. 2011. Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb argument processing. *Lang Cogn Process*. 26:1402-1456.
- Paczynski M, Kuperberg GR. 2012. Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *J Mem Lang*. 67:426-448.
- Press C, Kok P, Yon D. 2020. The perceptual prediction paradox. *Trends Cogn Sci*. 24:13-24.
- Roehm D, Bornkessel-Schlesewsky I, Rosler F, Schlewsky M. 2007. To predict or not to predict: influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience*. 19:1259-1274.

- Rugg MD. 1990. Event-related potentials dissociate repetition effects of high and low frequency words. *Memory and Cognition*. 18:367-379.
- Sassenhagen J, Fiebach CJ. 2019. Finding the P3 in the P600: Decoding shared neural mechanisms of responses to syntactic violations and oddball targets. *Neuroimage*. 200:425-436.
- Sassenhagen J, Schlesewsky M, Bornkessel-Schlesewsky I. 2014. The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain Lang*. 137:29-39.
- Schwanenflugel PJ, Lacount KL. 1988. Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning Memory and Cognition*. 14:344-354.
- Siedenberg R, Goodin DS, Aminoff MJ, Rowley HA, Roberts TPL. 1996. Comparison of late components in simultaneously recorded event-related electrical potentials and event-related magnetic fields. *Electroencephalogr Clin Neurophysiol*. 99:191-197.
- Steinhauser M, Yeung N. 2010. Decision processes in human performance monitoring. *J Neurosci*. 30:15643-15653.
- Taylor W. 1953. 'Cloze' procedure: A new tool for measuring readability. *Journal Q*. 30:415-433.
- van de Meerendonk N, Kolk HHJ, Chwilla DJ, Vissers CTWM. 2009. Monitoring in language perception. *Lang Linguist Compass*. 3:1211-1224.
- van den Brink D, Brown CM, Hagoort P. 2001. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *J Cogn Neurosci*. 13:967-985.
- Vespignani F, Canal P, Molinaro N, Fonda S, Cacciari C. 2010. Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*. 22:1682-1700.
- Young MP, Rugg MD. 1992. Word frequency and multiple repetition as determinants of the modulation of event-related potentials in a semantic classification task. *Psychophysiology*. 29:664-676.