

## RESEARCH ARTICLE

# The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG

William Matchin<sup>1</sup>  | Christian Brodbeck<sup>2</sup> | Christopher Hammerly<sup>3</sup> | Ellen Lau<sup>4</sup>

<sup>1</sup>Department of Communication Sciences and Disorders, University of South Carolina, Columbia, South Carolina

<sup>2</sup>Institute for Systems Research, University of Maryland, College Park, Maryland

<sup>3</sup>Department of Linguistics, University of Massachusetts, Amherst, Massachusetts

<sup>4</sup>Department of Linguistics, University of Maryland, College Park, Maryland

## Correspondence

William Matchin, Department of Communication Disorders, University of South Carolina, Discovery 1, Room 202D, Columbia, South Carolina 29208.

Email: matchin@mailbox.sc.edu

## Funding information

University of Maryland

## Abstract

Humans have a striking capacity to combine words into sentences that express new meanings. Previous research has identified key brain regions involved in this capacity, but little is known about the time course of activity in these regions, as hemodynamic methods such as fMRI provide little insight into temporal dynamics of neural activation. We performed an MEG experiment to elucidate the temporal dynamics of structure and content processing within four brain regions implicated by fMRI data from the same experiment: the temporo-parietal junction (TPJ), the posterior temporal lobe (PTL), the anterior temporal lobe (ATL), and the anterior inferior frontal gyrus (IFG). The TPJ showed increased activity for both structure and content near the end of the sentence, consistent with a role in incremental interpretation of event semantics. The PTL, a region not often associated with core aspects of syntax, showed a strong early effect of structure, consistent with predictive parsing models, and both structural and semantic context effects on function words. These results provide converging evidence that the PTL plays an important role in lexicalized syntactic processing. The ATL and IFG, regions traditionally associated with syntax, showed minimal effects of sentence structure. The ATL, PTL and IFG all showed effects of semantic content: increased activation for real words relative to nonwords. Our fMRI-guided MEG investigation therefore helps identify syntactic and semantic aspects of sentence comprehension in the brain in both spatial and temporal dimensions.

## KEYWORDS

fMRI, MEG, prediction, semantics, sentence processing, syntax

## 1 | INTRODUCTION

A central property of human language is our ability to combine words into sentences to express novel meanings (Chomsky, 1965; Jackendoff, 2002). In order to identify brain systems involved in combinatory processing, many neuroimaging studies using functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have utilized contrasts of structure, comparing highly structured stimuli such as sentences (e.g., *the poet will recite a verse*) to less structured stimuli such as word lists (e.g., *rabbit the could extract catch protect*). The key assumption is that brain areas engaged in syntactic and semantic combinatory processing will show increased activation to sentences, while noncombinatory effects of sensory, attentional, and lexical processing are subtracted out (Fedorenko, Nieto-Castanon, & Kanwisher, 2012; Humphries, Binder, Medler, & Liebenthal, 2006; Humphries, Love, Swinney, & Hickok, 2005; Matchin, Hammerly, & Lau, 2017; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene,

2011; Rogalsky & Hickok, 2008; Stowe et al., 1998; Vandenberghe, Nobre, & Price, 2002; see Zaccarella, Schell, & Friederici, 2017, for a meta-analysis and review). While there is some degree of variability across studies, four broad portions of the left hemisphere reliably show effects of sentence structure: the temporo-parietal junction (TPJ),<sup>1</sup> the anterior temporal lobe (ATL), the posterior temporal lobe (PTL), and the inferior frontal gyrus (IFG).<sup>2</sup> While fMRI and PET have provided precise spatial information about combinatory processing, these methods have poor temporal resolution, and so the time course

<sup>1</sup>We note that this is not a precise anatomical term. Some previous authors have described this region as the angular gyrus, but the relevant functional region encompasses a more anterior and inferior portion of the angular gyrus, including activations that extend beyond the angular gyrus into the middle temporal gyrus and supramarginal gyrus. Therefore, we use the term TPJ to be more inclusive and to fit with our ROI selection procedure.

<sup>2</sup>Here we describe the regions relevant for language quite coarsely; finer anatomical distinctions within these broad regions are clearly relevant. We take this broad approach because the spatial imprecision of MEG does not allow us to make finer anatomical claims about our effects.

of these structural effects remains largely unknown. The present study consisted of a within-subjects multimodal imaging approach using magnetoencephalography (MEG) to reveal the temporal dynamics of sentence processing within language-sensitive regions of interest (ROIs) defined by fMRI. The MEG signal measures magnetic fields induced by intracranial currents, thought to be due largely to synaptic activity at cortical pyramidal cells (Buzsáki, Anastassiou, & Koch, 2012). MEG thus reflects similar albeit not identical physiological processes as fMRI. Because magnetic fields propagate without temporal delay, brain activity can be measured with millisecond accuracy, complementing the excellent spatial resolution of fMRI (Hall, Robson, Morris, & Brookes, 2014). MEG can thus provide information about the timecourse of effects which previous work only localized spatially. The added temporal information allows conclusions about the functions of these regions with respect to syntactic and semantic processing not possible through the use of hemodynamic methods alone.

In natural language, combinatory semantic operations are tightly tied to syntactic structure. Distinguishing whether structural effects are due to syntax or semantics is possible using a contrast of content: comparing natural sentences, with all real words, to *jabberwocky* sentences (Carroll, 1871), in which the content words are replaced with nonwords while the function words are the same (e.g., *the tevill will sawl a pand*). This greatly reduces conceptual-semantic content while the syntactic structure is preserved. Interestingly, such contrasts typically reveal robust increased activation in structure-sensitive regions (Fedorenko et al., 2012; Humphries et al., 2006; Matchin et al., 2017), raising questions about the separability of syntax and semantics in the brain.

However, a separate body of work has associated the IFG and the PTL with lexical retrieval and lexical storage, respectively (see Lau, Phillips, & Poeppel, 2008 for a review). All syntactic theories posit some amount of stored syntactic structure associated with lexical items (Chomsky, 1981, 1995; Goldberg, 1995; Joshi & Schabes, 1997; Pollard & Sag, 1994), and stored structure is common in computational sentence processing models (Demberg, Keller, & Koller, 2013; Lewis & Vasishth, 2005; Vosse & Kempen, 2000). Thus, lexical and syntactic processing may be part of a single function of the PTL and the IFG, namely lexical-syntactic processing (see Snijders et al., 2008 for a similar proposal). On this view, effects of content in the PTL and the IFG could reflect the more robust activation of stored lexical entries (with associated syntactic structure) in the natural condition triggered by real content words, and effects of structure could reflect the activation of both lexicalized syntax as well as syntactic representations associated with sentence level structure unconnected to individual lexical items (such as clausal nodes with subject noun phrase and predicate verb phrase nodes; Jackendoff, 2017). By contrast, both the ATL and the TPJ are strongly associated with conceptual information (Binder, Desai, Graves, & Conant, 2009; Lau, Weber, Gramfort, Hämäläinen, & Kuperberg, 2014; Price, Bonner, Peelle, & Grossman, 2015), suggesting that effects of content in these regions reflects conceptual-semantic processing rather than lexical or syntactic processing. Supporting this hypothesized distinction, contrasts of *jabberwocky* sentences with unstructured *jabberwocky* lists (e.g., *his should woon pald the must*), reliably activates the IFG and the PTL but typically not the ATL and the TPJ (Fedorenko et al., 2012; Goucha &

Friederici, 2015; Matchin et al., 2017; Pallier et al., 2011), suggesting that IFG and PTL are involved in syntax regardless of the richness of conceptual-semantic content.<sup>3</sup>

A critical test for such hypotheses is the relative timing of activation across the course of a sentence. In particular, a substantial body of research has found that syntactic processing can be highly predictive (Levy & Keller, 2013; Omaki et al., 2015; Staub & Clifton, 2006; Sturt & Lombardo, 2005). If the IFG and the PTL are involved in lexical-syntactic processing, they might show early effects of structure. Previous work has suggested that the ATL is involved in local conceptual-semantic processing (Bemis & Pykkänen, 2011; Patterson et al., 2007; Westerlund, Kastner, Al Kaabi, & Pykkänen, 2015), and may be tied to the N400 component associated with lexical-conceptual processing in ERPs (Van Petten & Kutas, 1991; Lau et al., 2008; Kutas & Hillyard, 1983); this hypothesis would predict that the ATL would show content effects throughout the sentence. In contrast, recent work has suggested that the TPJ is more specifically involved in event semantics (Boylan, Trueswell, & Thompson-Schill, 2015, 2017; Grewe et al., 2007; Thompson et al., 2007), which would predict that the TPJ would show a late, post-verbal effect of content associated with the event semantics of natural sentences (which center around the verb). Although these distinctions are relatively coarse, obtaining initial data on the temporal dynamics of sentence processing within language-sensitive brain regions is a necessary first step toward the development of more detailed models of neurophysiological responses associated with combinatorial processing in the brain.

A few recent MEG (Brennan & Pykkänen, 2017; Halgren et al., 2002) and electrocortigraphy (ECoG) studies (Fedorenko et al., 2016; Nelson et al., 2017) have begun to address how the basic structural effects observed in fMRI and PET studies emerge over time. For instance, Nelson et al. (2017) found that the PTL (but not the IFG or ATL) shows a predictive response profile. By contrast, Fedorenko et al. (2016) found largely similar temporal dynamics across these regions. On the other hand, an MEG study by Brennan and Pykkänen (2017) found a significant fit of activity in the ATL with a mildly predictive sentence parser, and did not obtain a significant fit in the IFG. However, this study did not test for effects in the PTL and did not distinguish between syntactic and semantic processing in the ATL. The mixed results of these recent studies are likely due to differences in sentence materials, task, and neuroanatomical coverage. The goal of the current study was to more directly evaluate the timecourses of the fMRI structure and content effects by performing a parallel, within-subjects MEG and fMRI experiment using the same paradigm with different materials.

We used a 3 (STRUCTURE)  $\times$  2 (CONTENT) parametric design across both fMRI and MEG sessions. We investigated effects of conceptual-semantic CONTENT using *jabberwocky*, replacing open-class words with phonologically plausible nonwords. We investigated effects of STRUCTURE by contrasting three lexically matched levels of structure: six-word sentences, two-word phrases, and randomized word lists. We used a single sentence structure and a block design in order to

<sup>3</sup>Some studies have found *jabberwocky* structure effects in the ATL (Mazoyer et al., 1993; Humphries et al., 2006), but these effects are notably weaker and less consistent than in the IFG and the PTL.

encourage predictive syntactic processing and to limit the possibility that subjects would erroneously attempt to impose structure on unstructured conditions. Group fMRI data from this study have been previously reported (Matchin et al., 2017). In keeping with the literature, we found effects of *STRUCTURE* (sentences > phrases and lists) and *CONTENT* (natural sentences > jabberwocky sentences) in the IFG, PTL, ATL, and TPJ, but we did not find differences between the minimal two-word phrases and lists in the IFG, PTL, or TPJ.

In the current study, we use the MEG data from the same subjects to estimate the timecourse of the most robust *STRUCTURE* and *CONTENT* effects observed in fMRI. Prior work has argued that using ROIs derived from group data results in a significant loss of power due to individual variability in functional neuroanatomy (Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010; Fedorenko & Kanwisher, 2009; Rogalsky, Almeida, Sprouse, & Hickok, 2015). Therefore, we used the fMRI data at the subject level to generate individualized IFG, PTL, ATL, and TPJ ROIs for MEG data analysis. We recognize that analyzing MEG data within ROIs defined by fMRI data is potentially subject to error because MEG is less spatially precise than fMRI; therefore, we supported our ROI analyses with exploratory whole-brain analyses reported in Supplementary Materials.

We focus our presentation here on a smaller part of the larger design, choosing pairwise contrasts of *STRUCTURE* (six-word sentences vs. two-word phrases) and *CONTENT* (natural sentences vs. jabberwocky sentences). We did this for two reasons. First, the reduced design allows us to present clear and interpretable timecourses that focus on the maximal fMRI effects. Second, the word list conditions are problematic for MEG because open-class and closed-class words have markedly different neurophysiological response profiles (Halgren et al., 2002; Kutas & Hillyard, 1983; Münte et al., 2001; Van Petten & Kutas, 1991). This means that any contrast between conditions at a particular word position would be conflated with word class; given our use of dSPM noise normalization (Dale et al., 2000), selecting only trials where the open-closed distinction is matched across conditions and positions would greatly reduce statistical power. We designed our two-word phrase condition to solve this open/closed problem for MEG by matching the position of open-/closed-class items with the sentence condition within structurally unconnected phrases (Figure 1). In fMRI this is not a problem because the signal is integrated across the entire sentence. We therefore included the word list condition in both the fMRI

and MEG sessions for crucial comparison to the existing fMRI literature and to keep the experiment constant between sessions.

We expected the PTL and the IFG to show effects of *STRUCTURE* (greater activity for natural sentences relative to natural phrases) occurring on word 1 and/or word 2, given that these brain regions are associated with syntactic processing and previous evidence that the PTL shows a predictive response profile (Nelson et al., 2017). We expected that the ATL and the TPJ would show effects of *STRUCTURE* (greater activity for natural sentences relative to natural phrases) on word 5 and/or word 6, given that these regions are more associated with conceptual-semantic processing and previous evidence that the ATL shows a bottom-up response profile (Nelson et al., 2017; Bemis & Pykkänen, 2011). We expected that the ATL, PTL, and IFG would show overall main effects of *CONTENT* (increased activation for real words relative to nonwords), given a role for conceptual-semantic processing in the ATL, and for lexical processing in the PTL and the IFG. We expected that the TPJ might show effects of *CONTENT* (greater activity for natural sentences relative to jabberwocky sentences) at words 4, 5, or 6 (i.e., on or after the verb), given a possible role for this region in event semantics.

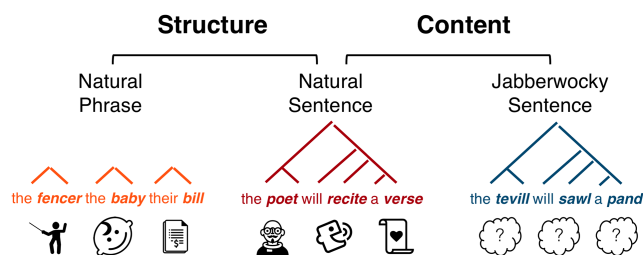
## 2 | MATERIALS AND METHODS

### 2.1 | Subjects

15 subjects were included in the analyses presented here (8 women) (mean age: 24, range: 20–29). One additional subject reported in Matchin et al. (2017) was excluded here due to a lack of significant activations in the fMRI experiment to support the individual ROI approach; one other subject was excluded from both reports due to excessive movement in the fMRI session. Additionally, 1 of the 15 remaining subjects only had significant fMRI effects in the IFG and TPJ—for this reason, analyses within the IFG and TPJ ROIs included 15 subjects, and analyses within the ATL and PTL ROIs included 14 subjects. Subjects were right-handed, native speakers of English, with normal or corrected-to-normal vision. Consent was acquired from each subject before the study began, separately for the fMRI and MEG sessions, and all procedures were approved by the Institutional Review Board of the University of Maryland. The order of fMRI and MEG sessions for the total 16 subjects was counterbalanced, with a minimum of 2 days in between sessions (mean: 12 days, standard deviation: 9 days).

### 2.2 | Stimuli and design

Our experiment used a 3 (*STRUCTURE*) × 2 (*CONTENT*) design, with each stimulus consisting of a six-word sequence. The three levels of structure were (i) LIST, (ii) PHRASE, and (iii) SENTENCE, and the two levels of content were (i) NATURAL and (ii) JABBERWOCKY. All of these conditions were presented in both the fMRI and MEG experiments. Here we report MEG analyses of a subset of this fuller design, three conditions: (i) three two-word phrases with real open-class or content words (*NATURAL PHRASE*), (ii) sentences with real open-class or content words (*NATURAL SENTENCE*), and (iii) sentences with content words replaced with nonwords (*JABBERWOCKY SENTENCE*). All conditions included real closed-class or function words. For the reasons described above, the MEG analyses reported in



**FIGURE 1** Schematic of stimulus design with examples. Tree diagrams represent the constituent structure in each condition. Images underneath words indicate conceptual-semantic representations denoted by real words while clouds represent the lack of conceptual-semantic information associated with nonwords. Bold and italicized words indicate open-class/content words [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

the main text focus on the NATURAL SENTENCE VERSUS NATURAL PHRASE contrast of sentence-level STRUCTURE, and the NATURAL SENTENCE VERSUS JABBERWOCKY SENTENCE contrast of conceptual–semantic CONTENT (Figure 1).

Stimulus creation and presentation is described in detail in Supporting Materials. To summarize, we created large sets of sentences with identical syntactic structure as depicted in Figure 1, middle: simple, active voice, using a transitive verb and a modal auxiliary. Both the subject and object always consisted of a determiner (article or possessive pronoun) and a noun. For the JABBERWOCKY SENTENCE condition, we modified the content words from the set of NATURAL SENTENCE stimuli using the Wuggy software (Keuleers & Brysbaert, 2010) to create nonwords that were pronounceable and had subsyllabic structure and phonological transitional probabilities that were similar to the real content words. We then scrambled the position of these non-content words across different sentences, keeping the positions of the function words fixed, in order to eliminate superficial phonological similarities between the JABBERWOCKY SENTENCE and NATURAL SENTENCE conditions. To create the NATURAL PHRASE condition, we pulled out pairs of function and content words from the NATURAL SENTENCE condition (e.g., Determiner-Noun and Modal-Verb), and created sequences that consisted of three isolated phrases of the same type: [Det-Noun Det-Noun Det-Noun] and [Modal-Verb Modal-Verb Modal-Verb]. We then grouped six stimuli from each condition together to form blocks (36 words total). Each NATURAL PHRASE block consisted of four six-word sequences of Det-Noun and two six-word sequences of Modal-Verb, with order of Det-Noun and Modal-Verb sequences randomized within blocks. We greatly minimized lexical repetition within stimulus materials and between the fMRI and MEG sessions for each subject, counterbalancing specific lexical items for each condition across subjects. All of our stimuli can be found in Appendix A of Matchin et al. (2017).

## 2.3 | Language probe task

Subjects were told that their task was to respond to single-word memory probes that were to appear after some but not all sequences (2 of 6 sequences in each block). These sequences were selected randomly with the stipulation that at least one of the last two sequences had a probe word—this ensured that subjects would not encounter two probes early in the block and then stop paying attention. Probes matched a word within the sequence 50% of the time. Probes from within the sequence were randomly chosen with equal frequency across different positions in the sequence, and “lure” probes not from the sequence were chosen at random from a separate list of open-class items that were never used in the sequences. We did not present closed-class lure probes because of the limited set of closed-class items in the experiment. Because closed-class items were frequently repeated, presenting new closed-class items might be too easy, and presenting closed-class items from within our set but not in the targeted sequence might be too difficult. However, we did include closed-class probes from within the sequence to ensure that participants would be motivated to attend to the whole sequence (and not just the content words). We asked subjects to press a button indicating whether or not they recognized the word from the immediately preceding sequence, and to make their response after the onset of

the probe and before the onset of fixation for the next sequence. In the MEG session, subjects made their “yes” responses with one hand and “no” responses with the other, counterbalanced across subjects. In the fMRI session, subjects made their yes/no responses on a button box in their left hand only to minimize task-related activations of the left hemisphere.

## 2.4 | Procedure

We presented the experiment visually using the PsychToolBox Matlab package (Brainard & Vision, 1997; Kleiner et al., 2007; Pelli, 1997). Stimuli were presented with rapid serial visual presentation (RSVP), each item (word or number) appearing on the screen for 350 ms, with 150 ms of blank screen following each word for a total of 500 ms per word. Each sequence of six words was preceded by a fixation cross which appeared on the screen for 1,250 ms, with 150 ms of blank screen following, in order to give the subjects time to blink in-between sequences. On probe trials (two out of six sequences per block; randomly selected), the last item of the sequence was followed by 200 ms of fixation, 100 ms of blank screen, and the probe word with a question mark (e.g., *should?*) which appeared on the screen for 1,700 ms. Each trial block was preceded by fixation for 1,350 ms and 150 ms of blank screen, then a cue to the condition in that block (ENGLISH SENTENCE, JABBERWOCKY SENTENCE, ENGLISH PHRASE, etc.), which appeared on the screen for 800 ms with 150 ms of blank screen. We included these cues such that subjects would be prepared for that condition and adapt their processing mode appropriately, for example, not mistakenly beginning to parse the unstructured condition as structured or vice versa. This cue was then followed by 350 ms of fixation and 150 ms of blank screen, after which began the first sequence of items. The total duration for one experimental block was 32 s. The fMRI experiment had a nearly identical presentation to the MEG experiment. In fMRI, the initial fixation preceding the onset of each block was slightly lengthened from 1,250 to 1,350 ms, and the fixation preceding each sequence of six words was shortened from 1,350 to 650 ms, since we did not need to provide a long period of time for subjects to blink. The MEG experiment consisted of 16 blocks from all 7 conditions (6 language conditions plus a nonlinguistic baseline MATH condition not reported here), for a total of 112 stimulus blocks, presented in 8 experimental runs of 14 blocks (2 blocks per condition, random order). The experiment lasted approximately 1.5 hr in total. The fMRI experiment was similar, except we presented 20 blocks from all of 8 conditions (6 language conditions, the math condition, and rest trials), for a total of 160 stimulus blocks, presented in 10 runs of 16 blocks each (2 blocks per condition, random order), lasting approximately 1.75 hr. The order of fMRI and MEG sessions was counterbalanced across subjects; sessions were separated by at least 2 days (mean: 12 days, standard deviation: 9 days).

## 2.5 | Behavioral data analysis

Here we report behavioral data while subjects performed the task inside the MEG scanner (we report behavioral performance during fMRI scanning in Matchin et al., 2017). We were unable to collect behavioral data from one subject due to technical issues, so we analyzed behavioral data from the remaining 14 subjects. We analyzed performance on



open-class items only, as open-class probes contained both items from within the stimulus and lure trials (items not within the stimulus), while closed-class probes were always within the stimulus (i.e., we never presented a closed-class item that wasn't in the preceding sequence). Thus we could only calculate  $d'$  values that correct for response bias for open class items.

## 2.6 | fMRI data collection and analysis

The primary focus of the current report is the MEG data. However, we used individual-subject regions of interest defined by the fMRI data to constrain the MEG analyses, and therefore we describe fMRI methods here as well. All fMRI methods, processing, and analysis were identical to Matchin et al. (2017) except for the use of individual-subject peak selection. During the fMRI experiment, MR images were obtained in a Siemens TRIO 3T scanner (Siemens Medical Systems) using a 32-channel head coil. We first collected a high-resolution T1-weighted anatomical image in the axial plane (voxel dimensions: 0.45 mm  $\times$  0.45 mm  $\times$  0.9 mm). We then collected a total of 2,370 T2\*-weighted EPI volumes over 10 runs. Each volume contained 36 oblique slices oriented approximately 20° clockwise relative to the AC-PC axis (TR = 2 s, TE = 25 ms, flip angle = 90°, in-plane resolution = 3 mm  $\times$  3 mm, slice thickness = 3 mm with 0.3 mm gap).

### 2.6.1 | Preprocessing

The first four volumes of each run were collected before stimulus presentation and discarded to control for T1 saturation effects. Slice-timing correction, motion correction, and spatial smoothing were performed using AFNI (Cox, 1996; <http://afni.nimh.nih.gov/afni>). Motion correction was achieved by using a 6-parameter rigid-body transformation, with each functional volume in each run first aligned to a single volume in that run. Functional volumes were aligned to the anatomical image, and subsequently aligned to Talairach space (Talairach & Tournoux, 1988). Functional images were resampled to 3 mm isotropic voxels and spatially smoothed using a Gaussian kernel of 6 mm FWHM. The data were high-pass filtered with a cutoff frequency of 0.0056 Hz at the first-level analysis stage using AFNI's 3dDeconvolve function using the "polort" parameter with a value of 4.

### 2.6.2 | Defining individual subject regions of interest

Rather than using anatomical ROIs, which do not account for individual subject anatomical and functional variability (particularly with respect to language), we generated subject-specific ROIs from the functional activation peaks in the fMRI data (Fedorenko & Kanwisher, 2009; Fedorenko et al., 2010; Nieto-Castañón & Fedorenko, 2012; Rogalsky et al., 2015). This ensured that the most functionally relevant regions in each subject were identified for MEG analysis within broadly the same anatomical regions across subjects.

Individual subject analyses were performed using AFNI's 3dDeconvolve function. The regression analysis was performed to find parameter estimates that best explained variability in the data. Each predictor variable representing the time course of stimulus presentation was entered into a convolution analysis using a canonical hemodynamic response function (AFNI's BLOCK parameter). The following seven regressors of interest were included in the model: NATURAL

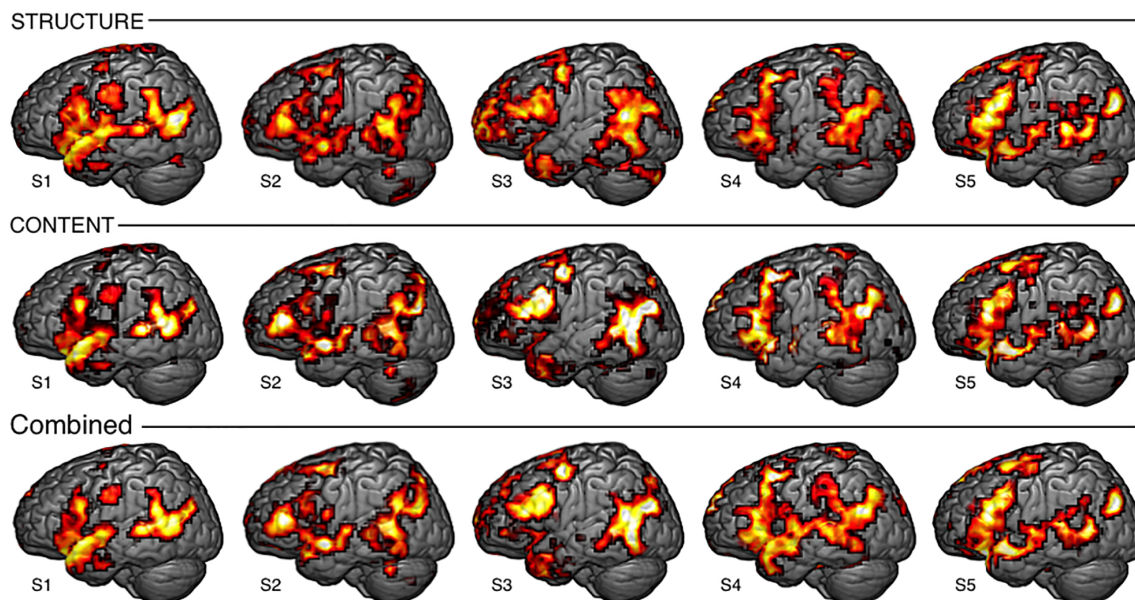
SENTENCE, JABBERWOCKY SENTENCE, NATURAL PHRASE, JABBERWOCKY PHRASE, NATURAL LIST, JABBERWOCKY LIST, and MATH. The six motion parameters were included as regressors of no interest.

The natural sentence > natural phrase and natural sentence > jabberwocky sentence effects highly overlapped at both the group and individual subject level, with nearly identical peak coordinates for these effects for most subjects (see Figure 2, rows 1 and 2, for statistical maps for five example subjects). Therefore, we combined the two contrasts by averaging the statistical maps produced by each contrast, which resulted in a single statistical map incorporating information from both contrasts (Figure 2, row 3). This allowed us to select the same set of ROIs for both contrasts in MEG. We chose to average statistical maps rather than taking the intersection so that we could select ROIs using peak  $t$ -values; without local maxima, it would be difficult to disentangle overlapping clusters. For each subject, we selected peak coordinates from the averaged activation maps, using a threshold of  $p < .005$  (one-tailed) and a cluster size correction of 20 voxels (a similar procedure to Rogalsky et al., 2015). We did this to select ROIs that were optimized for each subject by individual functional activation rather than for the purposes of statistical inference about these activation patterns and to identify coordinates for the largest amount of our subjects, although most subjects had significant activations using a much stricter statistical threshold (see Table 1 for  $t$ -values for each individual peak coordinate used in this study).

We selected four coordinates for each subject using the strongest local peak within four, manually determined pre-defined general anatomical regions (Figure 3, left): the TPJ, PTL, ATL, and IFG. The TPJ included peaks in the middle temporal gyrus (MTG), supramarginal gyrus, or angular gyrus posterior to the end of the sylvian fissure (peaks in MTG were prioritized). The PTL included peaks in the superior temporal sulcus (STS) or MTG, posterior to primary auditory cortex and anterior to the end of the sylvian fissure. The ATL included peaks in the STS or temporal pole anterior to primary auditory cortex (if peaks occurred in both temporal pole and STS, STS peaks were used). The IFG included peaks in the anterior portion of Broca's area, typically in the pars triangularis, with some peaks on the border of the pars triangularis and pars orbitalis. One subject did not have any significant clusters for the averaged statistical map, but did have clusters in the TPJ and IFG for the effect of content; therefore, we used these coordinates for that subject. This resulted in 15 subjects for the TPJ and IFG analyses, and 14 subjects for the ATL and PTL analyses. To display these coordinates, we first warped the subject's anatomical MRI to Talairach space using AFNI's @auto\_tlrc script, then applied the resulting transformation matrix to the coordinates in native space to generate coordinates in Talairach space. We then created a sphere with radius 5 mm around each coordinate and plotted all spheres on a 3D rendered template brain (Figure 3, right). The set of selected coordinates in Talairach space for all subjects and regions is listed in Table 1.

## 2.7 | MEG data collection and analysis

Prior to recording, five head position indicator coils were affixed to each subject's head, and the position of these coils, as well as the subject's headshape, were digitized using a Polhemus 3SPACE FASTRAK



**FIGURE 2** Left hemisphere statistical maps for the effect of structure, natural sentence > natural phrase (top), the effect of content, natural sentence > jabberwocky sentence (middle), and the combined average of those two effects (bottom) for five individual subjects, displayed on a template brain in Talairach space (Talairach & Tournoux, 1988) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

system. The indicator coils were used to determine the placement of the subject's head in the MEG dewar for source analysis (see *Source estimation* below). During the experimental sessions, subjects laid supine in a dark magnetically shielded room (Yokogawa Industries, Tokyo, Japan). Continuous MEG recording was executed using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan), and data was sampled at 500 Hz (60 Hz online notch filter, DC–200 Hz recording bandwidth).

### 2.7.1 | Preprocessing

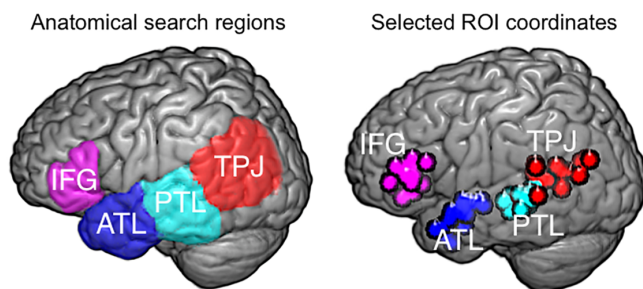
Environmental noise was removed from the data by regressing the signal at each channel onto three orthogonal reference channels located

far from the participant's head (Adachi, Shimogawara, Higuchi, Haruta, & Ochiai, 2001) using the system vendor provided recording software. Data were then analyzed with MNE-Python v. 0.14 (Gramfort et al., 2013) and Eelbrain v. 0.26 (Brodbeck, 2017). Bad channels were identified based on visual inspection of the raw data and excluded (between 0 and 3 per subject), and extraneous artifacts were removed using temporal signal space separation (Taulu & Simola, 2006). Next, Independent component analysis (ICA) was used to remove ocular and cardiac artifacts. ICA decomposition was performed on the entire session's MEG data for each subject. First, principal component analysis (PCA) was used to reduce the complexity of the data, retaining enough PCA components to explain 99% of the variance. Then, ICA decomposition was

**TABLE 1** Full set of individual subject peak coordinates from fMRI data for the combined effect of STRUCTURE and CONTENT used to create ROIs for MEG data analysis, and the corresponding t-value for each coordinate

TPJ				PTL				ATL				IFG			
x	y	z	t	x	y	z	t	x	y	z	t	x	y	z	t
−47	−60	8	8.50	−56	−37	6	8.72	−51	−9	−1	7.11	−57	16	23	5.49
−40	−70	22	6.43	−55	−43	7	8.02	−54	3	−10	9.44	−48	38	9	7.84
−47	−70	16	5.14	−63	−39	−1	4.90	−49	−14	0	6.49	−31	38	2	4.08
−59	−47	23	5.24	−44	−42	6	7.59	−53	2	−6	5.77	−39	26	1	7.38
−65	−49	8	4.57	−45	−44	8	4.22	−37	4	−28	3.92	−53	34	19	4.46
−43	−51	12	10.00	−55	−42	0	9.56	−45	−4	−16	5.41	−47	24	18	7.20
−36	−55	17	9.14	−45	−44	5	8.82	−50	−4	0	6.73	−50	27	18	6.86
−47	−57	21	7.07	−43	−29	−2	5.35	−52	6	−14	8.83	−50	25	23	8.04
−41	−44	20	4.12	−53	−28	1	5.67	−37	8	−24	5.28	−53	25	10	5.67
−46	−60	18	4.78	−51	−31	−5	4.90	−51	−4	−7	8.65	−39	29	1	3.59
−34	−61	15	9.42	−58	−43	3	5.97	−47	−15	−2	6.23	−47	20	8	7.69
−36	−73	23	10.04	−50	−35	2	10.64	−56	11	−16	8.50	−52	24	14	8.09
−45	−44	16	6.49	−58	−41	5	4.17	−52	−16	−2	5.47	−54	28	14	5.85
−39	−48	12	7.40	−48	−31	3	7.98	−51	−6	−8	7.86	−41	32	−2	7.64
−46	−76	25	3.52									−51	38	9	3.93

Note. Coordinates reported in Talairach space (Talairach & Tournoux, 1988). One subject lacked coordinates for the PTL and the ATL.



**FIGURE 3** Left: Manually determined anatomical search regions within which peak coordinates for each subject were selected. Right: Individual subject peak coordinates within these search regions for the combined effect of *STRUCTURE* and *CONTENT* used to create ROIs for MEG data analysis. Red: TPJ; cyan: PTL, blue: ATL, magenta: IFG [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

performed using the infomax method (Bell & Sejnowski, 1995). Artifactual components were identified by visual inspection of their topography and time course, resulting in the removal of between 2 and 5 components per subject. Data were then low-pass filtered at 20 Hz (FIR filter with MNE-Python default settings), because stimulus-locked electrical brain activity is typically observed below that frequency, and higher frequencies increase measurement noise. In addition, MEG device-specific properties are equivalent to a 0.16 Hz high pass filter. Data were epoched from  $-100$  to  $500$  ms relative to the onset of each word and downsampled by a factor of 2 to a sampling rate of 250 Hz. Epochs in which the signal exceeded a threshold value of 2 pT at any time were discarded. This criterion led to an exclusion rate per subject of at most 1.4% of trials. Condition averages were computed using the remaining good epochs. Because the blocked design of the experiment entailed that participants might already be anticipating condition-specific linguistic stimuli during the period prior to the first word, no baseline correction was applied. We used empty room data for computing the noise covariance matrix because of the potential issue that baseline periods likely contain cognitive activity involving semantic processes in our regions of interest (temporal lobe, temporal-parietal junction, IFG) (Binder et al., 1999, 2009; Binder, 2012; McKiernan, D'angelo, Kaufman, & Binder, 2006). Using baseline periods involving brain activity might have affected our ability to detect meaningful signals during the task that share similar patterns. Therefore, the sensor noise covariance was computed from 2 min of empty room data (without regularization) rather than using the pre-stimulus period.

### 2.7.2 | Source estimation

For each subject, a model of the cortical mantle was reconstructed from the structural MRI using FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl et al., 1999). The structural MRI was then coregistered with the head shape that had been recorded prior to MEG data collection. The two models were first aligned using the position of the nasion, and rotation parameters were determined using an iterative least squares algorithm minimizing the distance of the left and right tragus between MRI and head shape. This fit was manually adjusted based on the head shape points where appropriate. A source space was defined on the white matter surface based on a four-fold subdivision of an icosahedron, resulting in 2,562 sources in each hemisphere

with approximately even spacing. A forward solution was computed including three orthogonal dipoles for each source, one perpendicular and two parallel to the white matter surface. Minimum norm source estimates were computed with a depth weighting parameter of 0.8 (Lin, Witzel, et al., 2006), regularization  $\lambda = 1/9$ , loose orientation constraint (weighting factor 0.2; Lin, Belliveau, Dale, & Hämäläinen, 2006) and dSPM noise-normalization (Dale et al., 2000). The main advantage of dSPM is that spatial resolution is improved, with the complication that changes to the number of trials affects the noise estimates, and therefore the dSPM value. ROIs were constructed on the white matter surface source space, including all sources falling within a geodesic distance of 25 mm of the source closest to the corresponding fMRI peak. Averaged evoked fields for each condition were source localized, and average ROI time courses were extracted for group level analysis. For the purpose of whole sentence analysis and visualization, the whole sentence time course was reconstructed by concatenating source estimates from the responses to individual words (the 100 ms baseline period was discarded for all but the sentence-initial word).

### 2.7.3 | Statistical analysis

We performed two sets of ROI analyses: sentence-level and word-level. Sentence-level analyses were designed to identify potentially unique time periods exhibiting effects of structure or content tied to particular parts of the sentence and which might not be tightly aligned with word boundaries. These analyses examined our key contrasts of structure (natural sentence > natural phrase) and lexical content (natural sentence > jabberwocky sentence) across the entire 3,000 ms time period associated with each six-word trial. Word-level analyses were designed to incorporate into the statistical model any systematic impacts of structure and content in the response to each individual word across positions, taking into account known differences between the event-related response to open- and closed-class words. Here we conducted  $2 \times 3$  ANOVAs of *STRUCTURE* (natural phrase, natural sentence)  $\times$  *POSITION* (early, middle, late), and *CONTENT* (jabberwocky sentence, natural sentence)  $\times$  *POSITION* (early, middle, late) on the 500 ms time period associated with each word of the sequence. For the factor of *POSITION*, we only report interactions with *STRUCTURE* and *CONTENT*, as main effects of *POSITION* (i.e., effects that do not depend on either the *STRUCTURE* or *CONTENT* factors) are difficult to interpret and might be driven by nonlinguistic factors. Open- and closed-class items were analyzed separately due to their systematic differences in position (closed always preceding open), as well as prior work demonstrating substantially different neurophysiological responses for these two word classes, whether due to visual differences in word length, word frequency, or grammatical function (Halgren et al., 2002; Kutas & Hillyard, 1983; Van Petten & Kutas, 1991).

All ROI analyses used temporal cluster-based permutation tests to determine statistical significance with strong control over type I error (Nichols & Holmes, 2002). For each contrast, *t*- or *F*-values were computed separately for each time point. The resulting time courses were then thresholded at a value equivalent to an uncorrected *p*-value of .05. Clusters were formed based on adjacency in time, and the exceedance mass (sum of all *t*- or *F*-values in the cluster) was computed as the cluster statistic for each cluster. This procedure was

**TABLE 2** Significant effects of STRUCTURE OF STRUCTURE × POSITION interactions for word-level ROI analyses

	Closed-class items		Open-class items	
	Structure	Structure × position	Structure	Structure × position
TPJ	None	180–300 ms, $p = .0002$ 320–352 ms, $p = .0341$	472–500 ms, $p = .0343$	None
PTL	92–148 ms, $p = .0297$	None	284–332 ms, $p = .0013$	420–472 ms, $p = .0083$
ATL	400–448 ms, $p = .0321$	None	None	None
IFG	None	None	None	None

Note. All reported effects of STRUCTURE are positive, reflecting increased activation for NATURAL SENTENCE relative to NATURAL PHRASE.

repeated under 10,000 random permutations of the data to generate distributions for the largest exceedance mass value for each effect. In each permutation, the data were shuffled by switching condition labels within subject, and the largest exceedance mass value was retained for each  $t$ - or  $F$ -map. Finally,  $p$ -values were computed for the clusters found in the original data based on the proportion of permutations that yielded a larger exceedance mass.

Although our focus in this report is on determining the MEG time course of our structure and content contrasts in regions that elicited fMRI effects, in Supporting Materials we provide exploratory whole-brain analyses that provide the opportunity to identify relevant regions outside of the main language-related regions, as well as exploratory analyses of the full  $3 \times 2$  design in spite of the open-/closed-class confounds that this introduces in event-related MEG.

### 3 | RESULTS

#### 3.1 | Behavioral data

Subjects performed well on the task overall in the MEG session, with average  $d'$  values in all conditions (open-class items only) greater than 2.5 (by convention, a  $d'$  value of 1 is considered good performance): NATURAL SENTENCE: 4.319, NATURAL PHRASE: 4.929, NATURAL LIST: 3.444, JABBERWOCKY SENTENCE: 3.314, JABBERWOCKY PHRASE: 3.453, JABBERWOCKY LIST: 2.690. A  $3 \times 2$  ANOVA revealed a main effect of structure:  $F(2,26) = 8.250$ ,  $p = .002$ , a main effect of content:  $F(1,13) = 9.230$ ,  $p = .010$ , and no interaction:  $F(2,26) = 1.102$ ,  $p = .347$ . Post-hoc tests of structure revealed that subjects performed significantly better on sentences relative to lists:  $t = 2.894$ ,  $p = .022$ , better on phrases relative to lists:  $t = 3.960$ ,  $p = .001$ , and no difference between sentences and phrases:  $t = -1.495$ ,  $p = .440$  (Bonferroni corrected). These results indicate facilitatory effects of both structure and content, in line with previous studies (Brener, 1940; Marks & Miller, 1964; Miller & Isard, 1963; Miller et al., 1951) as well behavioral performance in the fMRI session

of this experiment (Matchin et al., 2017), suggesting that subjects indeed parsed the structured material and processed the content of the words.

#### 3.2 | MEG: Structure

Results of the statistical analyses are displayed in Tables 2–4. Time course plots illustrating effects of structure are shown in Figure 4, and time course plots illustrating effects of content are shown in Figure 5.

We observed several distinct effects of structure in the MEG time course for the posterior temporal (PTL) ROI, beginning early in the sentence but continuing throughout. Most notable was a large increase in the response to the first open-class item (word 2; subject noun) in the sentence condition vs. the phrase condition, with responses appearing to diverge at about 250 ms after the onset of the word. This difference appeared in the sentence-level analysis as a significant difference between 772 and 984 ms (272–484 ms after word 2 onset), and in the open-class word-level analysis as a significant structure × position interaction in open-class words between 420 and 472 ms; a smaller main effect of structure was also observed across all open-class positions between 284 and 332 ms. Post-hoc tests within the time window of the structure × position interaction revealed significantly increased activity for the sentence condition relative to the phrase condition at word 2 ( $t[14] = 2.35$ ,  $p = .034$ ), but no differences at words 4 and 6 ( $t[14] \leq 0.13$ ,  $p \geq .902$ ). There was also a significant early main effect of structure across all closed-class positions between 92 and 148 ms post-word-onset. As we discuss in more detail below, the fact that these structural effects were observed as early as the first phrase of the sequence (which was largely matched across sentence and phrase conditions) appears consistent with a role for the PTL in predictive syntactic computations. Additionally, in the sentence-level analysis we observed a significant increased response to sentences at word 5 (the determiner beginning the final noun phrase) between 2,180 and 2,284 ms (180–284 ms after word 5 onset), which mirrored a similar pattern in the TPJ and

**TABLE 3** Significant effects of CONTENT OF CONTENT × POSITION interactions for word-level ROI analyses

	Closed-class items		Open-class items	
	Content	Content × position	Content	Content × position
TPJ	None	None	None	None
PTL	208–272 ms, $p = .0062$	None	228–332 ms, $p = .0009$	None
ATL	None	None	224–348 ms, $p = .0032$	None
IFG	None	None	216–396 ms, $p = .0004$	None

Note. All reported effects of CONTENT are positive, reflecting increased activation for NATURAL SENTENCE relative to JABBERWOCKY SENTENCE.



**TABLE 4** Significant effects in the sentence-level ROI analyses

	Structure	Content
TPJ	1964–2096 ms, (word 4–5) $p = .0028$ 2,156–2,356 ms (word 5), $p = .0005$	2,200–2,280 ms (word 5), $p = .0324$
PTL	772–984 ms (word 2), $p < .0001$ 2,180–2,284 ms (word 5), $p = .0104$	1,200–1,272 ms (word 3), $p = .0306$ 1,724–1808 ms (word 4), $p = .0071$ 2,724–2,844 ms (word 6), $p = .0039$
ATL	None	None
IFG	None	796–896 ms (word 2), $p = .0187$ 1,204–1,292 ms (word 3), $p = .0421$ 1,728–1800 ms (word 4), $p = .0411$

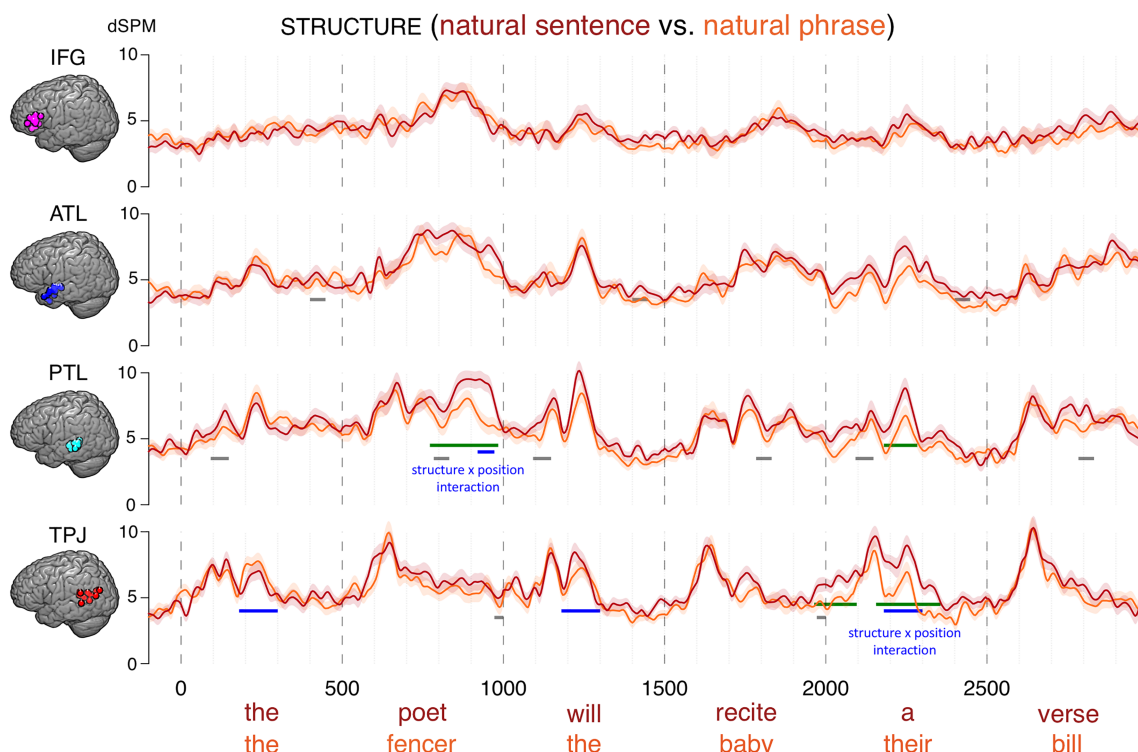
Note. All reported effects are positive, reflecting increased activation for NATURAL SENTENCE relative to NATURAL PHRASE (effects of STRUCTURE) or increased activation for NATURAL SENTENCE relative to JABBERWOCKY SENTENCE (effects of CONTENT).

the ATL that may reflect interpretive processes triggered by the verb phrase.

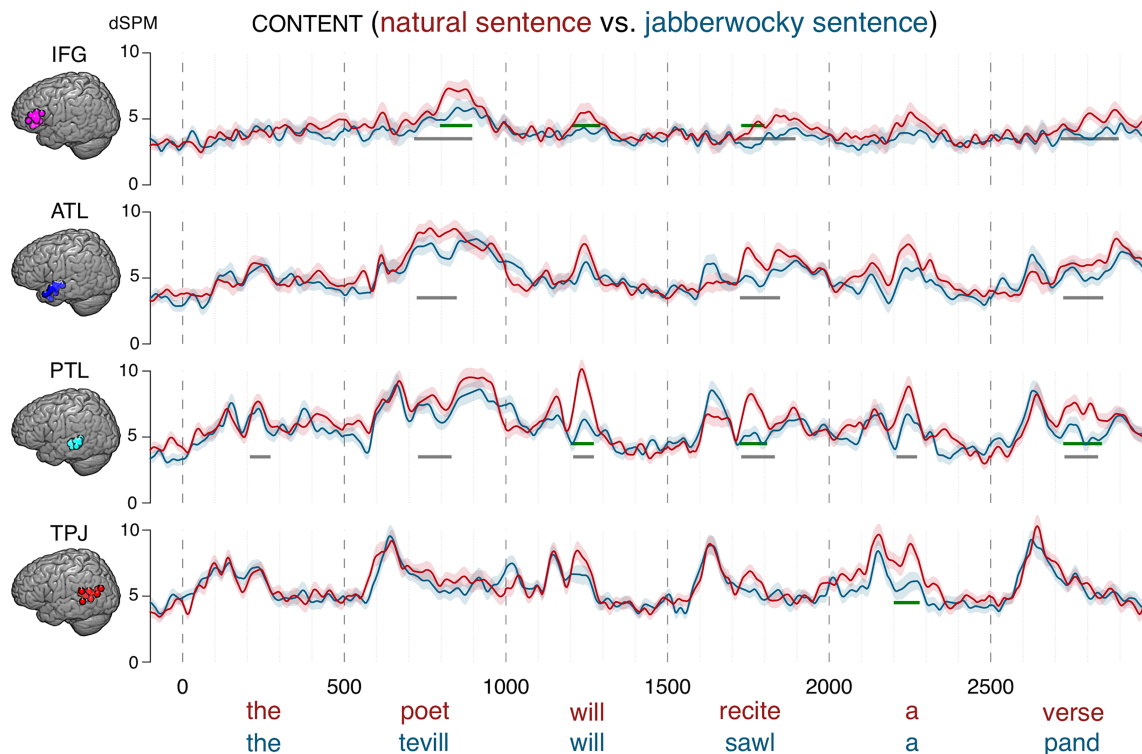
In the TPJ, this increased response to sentences vs. phrases at the onset of the last function word (word 5) was the largest structure effect, resulting in two nearly adjacent effects in the sentence analysis (between 1964–2096 ms and 2,156–2,356 ms, or from 36 ms prior to word 5 onset to 356 ms post-onset). We also observed a significant structure  $\times$  position interaction in the closed class word analysis that appeared to be due to the strong word 5 effect ( $t[14] = 3.97$ ,  $p = .001$ ), a similar but smaller effect at word 3 ( $t[14] = 2.54$ ,  $p = .024$ ), and a pattern in the reverse direction at word 1 ( $t[14] = 2.31$ ,  $p = .036$ ).

In the ATL we observed only a main effect of structure for closed-class words between 400 and 448 ms. In the sentence time courses, we note the presence of a nonsignificant but numerically sustained effect of structure between 2,000 and 2,500 ms (word 5) that parallels the effects observed in PTL and TPJ in the same time-window. Additionally, there appeared to be some short and sub-threshold effects of structure at the end of word 2.

In the IFG, there were surprisingly no effects of structure across all of our analyses. Examining the time courses for this ROI, there were only some weak and intermittent periods of time where natural sentences showed numerically increased activation relative to natural



**FIGURE 4** Analyses of STRUCTURE within each ROI (dSPM). Red: NATURAL SENTENCE, orange: NATURAL PHRASE. Gray lines indicate significant main effects of STRUCTURE in the word-level analyses, blue lines indicate significant interactions of STRUCTURE and POSITION in the word-level analyses, and green lines represent significant effects of STRUCTURE in the sentence-level analyses. X axis is time in milliseconds relative to onset of the first word in each six-word trial. Shading indicates the within-subject standard error (Loftus & Masson, 1994). The blue text “STRUCTURE  $\times$  POSITION interaction” and the corresponding blue underline reflects significant time periods for the STRUCTURE  $\times$  POSITION interaction in the word-level analysis. In the PTL, only the time period at word 2 survived a post-hoc pairwise comparison, while in the TPJ, the time periods at words 1, 3, and 5 all survived a post-hoc pairwise comparisons [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 5** Analyses of CONTENT within each ROI (dSPM). Red: NATURAL SENTENCE, blue: JABBERWOcky SENTENCE. Gray lines indicate significant main effects of CONTENT in the word-level analyses, and green lines represent significant effects of content in the sentence-level analyses. X axis is time in milliseconds relative to onset of the first word in each six-word trial. Shading indicates the within-subject standard error (Loftus & Masson, 1994) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

phrases. This was unexpected, given that we observed robust structure effects in the IFG in our fMRI data (Matchin et al., 2017).

### 3.3 | MEG: Content

While the timing of structure effects varied considerably across the four ROIs, the timing of lexical content effects appeared more uniform, and included the IFG. The IFG, ATL, and PTL all showed significantly greater activity for natural versus jaberwocky materials in the open-class words, roughly 215–350 ms post-word-onset. The PTL also showed a main effect of content for closed-class items, but interestingly this occurred in a later time window (208–272 ms) than the main effect of structure for closed-class items reported above (92–148 ms) and appeared especially large at the auxiliary verb position (word 3), suggesting distinct effects of structure and content on PTL responses to closed-class items. The IFG also showed a similar content effect at closed-class word 3 in the sentence-level analysis (204–292 ms post-word-onset). The only significant effect of content in the TPJ was in the sentence-level analysis at word 5 (200–280 ms post-word-onset), with approximately the same timing as the structure effect in this region.

## 4 | DISCUSSION

We used parallel fMRI and MEG to determine the time course of structural and conceptual-semantic processing in regions shown to be involved in combinatorial processing in previous fMRI and PET studies

(Fedorenko et al., 2012; Humphries et al., 2005, 2006; Matchin et al., 2017; Mazoyer et al., 1993; Pallier et al., 2011; Rogalsky & Hickok, 2008; Rogalsky et al., 2015; Stowe et al., 1998; Vandenberghe et al., 2002). Previous research in neurophysiology has examined general effects of sentence position, such as words early versus late in a sentence (Halgren et al., 2002; Fedorenko et al., 2016) or structural variables like parsing steps or constituent structure complexity (Brennan & Pykkänen, 2012; Nelson et al., 2017). Our results complement these studies by identifying specific points in a sentence when enhanced activity is seen for sentence structure and conceptual-semantic content. Given that our experiment was designed to maximize any effects of structural prediction beyond what is likely to occur in natural sentence processing, we do not claim that these temporal dynamics precisely characterize natural sentence processing under all conditions. Rather, the observed temporal dynamics help to characterize the underlying functions of these brain regions. Here we discuss the effects found in each ROI and their implications with respect to their functional contributions to sentence processing.

### 4.1 | The posterior temporal lobe: Lexical-syntactic processing

In previous work, we have hypothesized that the PTL underlies lexical-syntactic processing (Matchin, 2017; Matchin et al., 2017), and we have suggested two explanations for effects of structure in this region in fMRI and PET studies: (i) predictive activation of sentence-level syntactic representations, and (ii) increased attention or

maintenance of syntactic representations associated with lexical items when they are presented in a sentence context. The timing of the structure effects we observed here in MEG in our PTL ROI (based on fMRI activations in or near posterior superior temporal sulcus in individual subjects) appear to support both of these possibilities.

The early structure effect we observed in MEG at the end of word 2 (in the sentence-level analysis, 272–484 ms after onset of the subject noun) provides more direct evidence for the hypothesis that posterior temporal structure effects in part reflect syntactic predictions. This effect was reinforced by the cluster identified for this contrast in the supplementary whole brain analyses in a similar posterior temporal region, centered on the STS/MTG (Supporting Information Figure S3). Given the high similarity between the sentence and phrase conditions at this early stage in the sentence and our blocked experimental design that encouraged structural predictions, a straightforward interpretation of this effect is that in the sentence condition subjects processed the subject noun phrase and then projected sentence-level hierarchical structure before the appearance of the following words that support that structure. In the phrase condition, when subjects have clear information that there is no hierarchical sentence structure, subjects likely did not generate such structural predictions. This account naturally fits with predictive left-corner parsing models (Demberg et al., 2013; Lewis & Vasishth, 2005) and psycholinguistic data supporting structural predictions during sentence processing (Levy & Keller, 2013; Staub & Clifton, 2006; Sturt & Lombardo, 2005). These results also converge with those of Nelson et al. (2017), who found that activity in the PTL was best fit with a predictive top-down parsing model. Combined, these data suggest that sentence-level predictive structural processing can be localized at least in part to posterior regions of the temporal lobe, and more specifically that the head noun of the subject noun phrase may be a critical position in the sentence for projecting upcoming structure. Importantly, we do not claim that structural predictions always unfold with identical temporal dynamics as observed in this study, as our choice of sentence structure and block design likely substantially enhanced structural predictions. Rather, we suggest that the PTL encodes structural representations that can be used in a predictive fashion, particularly when context encourages the use of such predictions.

One potential caveat to note is that the distribution of words across positions of the sentence was not perfectly controlled across the sentence and phrase conditions. In the sentence condition, modal auxiliaries and verbs always occurred in the middle of each stimulus and determiners and nouns occurred at the beginning and ends of sentences, while in the phrase condition, these words were equally distributed throughout the stimulus. In principle, this might have affected our results, as a determiner + noun occurred 100% of the time early in the sentence stimuli, while a determiner + noun occurred only 67% of the time early in the phrase stimuli. However, we note that the middle and ends of stimuli are similarly imbalanced—this would predict an effect in the opposite direction at word 4, and a similar effect at word 6. However, this is not the pattern we see—there is in fact a small effect in the same direction at word 4, and no consistent effect at word 6 (there is a small effect in the same direction 228–332 ms after word onset, but the effect later reverses, ~400–500 ms after word onset). Thus, our results cannot be

straightforwardly attributed to differences in the distribution of words between the two conditions, but are transparently accounted for by higher-level structural differences.

The PTL showed two additional effects of structure that were distributed throughout the duration of the sentence: increased activation for sentences relative to phrases at closed-class words (words 1, 3, and 5) between 92 and 148 ms after word onset and at open-class words (word 2, 4, and 6) between 284 and 332 ms after word onset. Given that these effects occurred at all positions of the sentence, these effects could be explained by increased attention or maintenance of individual words with associated syntactic structure during sentence processing (Snijders et al., 2008). Under this hypothesis, during both the phrase and sentence conditions, words activate corresponding lexical–syntactic representations. In both conditions, this creates local, connected phrases. However, in the sentence condition, these local phrases are also integrated to create a higher-level sentence structure. This process would involve increased attention to individual syntactic representations, resulting in increased activation for these words in sentence context. The difference in the latencies of the structure effect in closed- and open-class words may simply reflect the fact that shorter and more frequent closed-class words can be identified more quickly.

We hypothesized that effects of content—that is, increased activation for natural sentences relative to jabberwocky sentences—in the PTL in fMRI and PET studies also reflect lexical–syntactic processing. Under this hypothesis, real words result in the selection of lexical–syntactic representations stored in the PTL while nonwords do not, resulting in greater overall lexical–syntactic processing in the PTL for natural sentences that have both real content and function words relative to jabberwocky sentences that only have real function words. Consistent with this hypothesis, we observed greater activation for real content words in natural sentences relative to nonwords in jabberwocky sentences in the PTL. This content effect highly overlapped in time with the structure effect for content words in all sentence positions (roughly 250–330 ms post-word-onset), suggesting that both of these effects may in fact be due to a common underlying source such as lexical–syntactic processing.

We also found greater activation for function words in natural sentences relative to jabberwocky sentences. Interestingly, while the effects of structure and content for content words highly overlapped in time, the effects on function words did not: for structure, the effect occurred between 92 and 148 ms after word onset; for content, the effect occurred between 208 and 272 ms. Visual inspection of the time courses shows a hint of a content effect for function words in the earlier time window at word 3, but no early effect at words 1 and 5, indicating that this lack of timing overlap was not an artifact of the temporal clustering procedure.

Here we offer a tentative hypothesis of the distinct timing of structure and content effects on function words. As discussed above, the effect of structure at function words between 92 and 148 ms may reflect increased attention to or maintenance of lexical–syntactic representations associated with those words in sentences relative to unconnected phrases. By contrast, the later effect of content may reflect stronger predictions of upcoming lexical items triggered by

function words in natural sentences relative to jabberwocky sentences. One important property of function words is that they select for particular phrasal complements (e.g., determiners select noun phrase complements, modal auxiliaries select verb phrase complements). This selection information is naturally useful to predict upcoming lexical items, since these function words must be followed by a phrase of a certain syntactic category. Such predictions of upcoming words would be equal in the natural sentence and natural phrase conditions, given that they both have the same set of local phrases with the same words (in other words, determiners are always followed by nouns and modals always followed by verbs in both conditions as well as the real world). However, in the jabberwocky sentence condition, function words are always followed by nonwords. It may be the case that lexical predictions triggered by function words are less robust or absent in jabberwocky sentences because such predictions never match a subsequent word. Supporting this interpretation, a previous MEG study showed that prediction of lexical items produced activation in the vicinity of this region (Dikker & Pykkänen, 2013). However, this hypothesis is tentative, and should be confirmed by a study more explicitly examining this issue.

Finally, we found a “late” structure effect in the PTL (natural sentences > natural phrases) in the sentence-level analyses at word 5 (180–224 ms after word onset). A similar effect was also found, more robustly, in the TPJ, and there appears to be a trend toward such an effect in the ATL. This effect appears to be a generally widespread increase in activation in all language regions for natural sentences relative to the other conditions (with the exception of the IFG) near the end of the sentence, as it can also be observed in the contrast between natural and jabberwocky sentences (although we did not find a significant sentence-level content effect at word 5 for the PTL). Other studies have also observed robust structural effects near the end of the sentence (Fedorenko et al., 2016; Nelson et al., 2017). As we discuss below in the section on the TPJ, we suggest that this effect is tied to the interpretation of the event semantics of the sentence, which might in turn drive additional lexical–syntactic processing in the PTL.

Interestingly, structure effects that we observed occurred at specific positions of the sentence, and there was no general increase in activation throughout the duration of the sentence, as hypothesized in Pallier et al. (2011) and observed by Fedorenko et al. (2016). This suggests that incremental sentence processing is not always supported by monotonic increased activity, but rather that particular syntactic and semantic operations are triggered at specific structural positions of a sentence. One explanation of the discrepancies between studies is that Fedorenko et al. (2016) may have collapsed activity across different sentence structures, producing an average general increase at later positions of a sentence that was driven by more punctate effects occurring at different parts of these sentences. These differences could also be due to the fact that we analyzed source activation magnitudes and Fedorenko et al. (2016) analyzed high-gamma power—it is possible that these different signals yield insight into complementary mechanisms underlying sentence processing.

## 4.2 | The anterior temporal lobe: Conceptual–semantic processing

The ATL ROI showed two significant effects in this study: increased activation for real content words in natural sentences relative to nonwords in jabberwocky sentences between 224 and 348 ms, and increased activation for identical function words in natural sentences relative to natural phrases between 400 and 448 ms. The increased activation for real words relative to nonwords is consistent with a role for this region in general conceptual–semantic processing (Nobre & McCarthy, 1995; Lau, Gramfort, Hämäläinen, & Kuperberg, 2013; Lau et al., 2014), and occurs in roughly the same time window associated with previous effects of conceptual–semantic combination in the ATL in MEG studies (Bemis & Pykkänen, 2011; 2012; Westerlund et al., 2015).

The late main effect of structure at function words potentially reflects increased attention to conceptual–semantic representations in sentence relative to phrase contexts. Because this effect occurred on function words that have much less rich conceptual–semantic content compared with content words, this effect may specifically reflect more robust predictions from function words to upcoming conceptual–semantic content in sentence contexts. The relatively weak effects of structure in this region may appear to conflict with the generally consistent finding of structural effects in the ATL in prior literature (Fedorenko et al., 2012; Humphries et al., 2005, 2006; Mazoyer et al., 1993; Pallier et al., 2011; Rogalsky & Hickok, 2008). However, this may reflect the fact that even the minimal syntactic/semantic combination in the two-word phrases acts to increase conceptual–semantic processing over and above the completely unstructured lists that are traditionally used: in our fMRI data (Matchin et al., 2017), the ATL ROIs (anterior STS and temporal pole) showed a more graded effect of structure than the PTL, with low activity for lists, higher activity for phrases, and strongest activity for sentences. Overall, our results are compatible with a role for the ATL in general conceptual–semantic processing, consistent with previous MEG studies (Halgren et al., 2002; Westerlund & Pykkänen, 2014).

## 4.3 | Temporo-parietal junction: Sentence-level conceptual–semantic processing

Several effects in the TPJ occurred near the end of word 4 and throughout word 5, close to the end of the sentence—increased activity at the end of natural sentences relative to jabberwocky sentences and natural phrases (this latter effect of structure was also clearly identified in the supplementary whole brain analyses in the TPJ, Supporting Information Figure S3). However, the TPJ did not show a significant effect of content for real content words in natural sentences relative to nonwords in jabberwocky sentences. This was somewhat surprising, given that all three other ROIs showed robust effects for this contrast in roughly the same time window, and that several previous studies have implicated the TPJ in word or phrase-level semantic processing (Binder et al., 1997, 2003; Rissman, Eliassen, & Blumstein, 2003; Bemis & Pykkänen, 2012; Price et al., 2015). However, many other studies have implicated the TPJ in argument structure and/or event-level semantics rather than word-level semantics (Boylan et al.,



2015; Grewe et al., 2007; Kalénine et al., 2009; Pallier et al., 2011; Thompson et al., 2007; Thothathiri, Kimberg, & Schwartz, 2012). If TPJ is primarily involved in processing coherent sentence meaning, its activity might not be tied to the specific conceptual content associated with individual lexical items. Our data are consistent with this view, as the increased activation for natural sentences near the end of the sentence suggests a response associated with the interpretation of the event semantics of the sentence, which is more complete at the end of the sentence than near the beginning.

Interestingly, this effect occurred *before* the final noun of the sentence. All of the verbs in our study were transitive, which means that the event's meaning is incomplete without two arguments (event participants). For instance, the meaning of an event associated with a verb like *destroy* is incomplete unless there is an agent of destruction and a recipient of destruction. One possibility is that our results reflect incremental interpretation of the sentence meaning, using only the subject and verb. For instance, in the sentence *a telepath could destroy our foe*, subjects could integrate *telepath* and *destroy* together before waiting for the object of the verb.

The follow-up pairwise tests to the structure  $\times$  position interaction showed that there was in fact increased activity for the phrase condition relative to the sentence condition at word 1, a function word (180–300 ms after word onset). It is unclear why phrases would activate the TPJ more than sentences at any time point, let alone at the beginning of the trial, although we note a similar but not significant effect in the PTL and ATL ROIs as well. The ECoG studies by Nelson et al. (2017, supporting information materials) and Fedorenko et al. (2016, supporting information materials) both report some indication of increased activation for word lists relative to sentences at early positions, suggesting that decreased activation in language regions for sentences relative to less structured conditions early in the trial is a general phenomenon and not an anomaly. This effect may be due to overall attentional differences between the two conditions, whereby subjects direct increased attention to sensory processing in the sentence condition early on, resulting in increased activation in systems involved in linguistic processing for the phrase condition. Future studies could attempt to narrow down more definitively the source of this early decreased activation for linguistic structure.

#### 4.4 | The inferior frontal gyrus: Limited effects

Our parallel fMRI study identified strong effects of structure and content in the left anterior IFG, replicating prior fMRI findings (Fedorenko et al., 2012; Humphries et al., 2006). In MEG we obtained robust effects of content in IFG for real content words relative to nonwords in a similar time window as the PTL and ATL, but we did not observe any effects of structure in this region, similar to several previous MEG studies (Bemis & Pykkänen, 2011; 2012; Brennan & Pykkänen, 2012; 2017; Westerlund & Pykkänen, 2014). The lack of these structure effects limits our ability to make inferences about the function of IFG from this study. One factor that could have contributed is reduced signal from frontal regions, as in standard MEG systems the subject's head rests against the back of the helmet, and most of the gap between the head size and the helmet size is thus in the frontal areas

(Marinkovic, Cox, Reid, & Halgren, 2004). Second, the cortical morphology of the IFG might be such that it is difficult to identify the relevant functional sources using MEG (Ahlfors et al., 2010), suggesting that future similar studies might profit from using electrophysiological methods, which have complementary orientation sensitivity. However, the fact that we still obtained robust effects of our content manipulation in IFG casts some doubt on this explanation, especially given that IFG effects of content and structure were relatively comparable in magnitude in our fMRI study. Other possible explanations relate to the different temporal properties of fMRI and MEG. If IFG effects of structure occur at variable timepoints from trial to trial, they would sum across trials in temporally blurred fMRI analyses, but would not sum across trials in temporally sensitive MEG analyses. Or, IFG effects of structure might primarily reflect “wrap-up” activity after the sentence is over; we were unable to evaluate this possibility in the current MEG dataset because of muscle movement associated with the probe response and increased ocular artifact post-sentence. Finally, although both fMRI and MEG signals are correlated with the local field potential, the relationship between the two measures is a complex one (Hall et al., 2014), and it is possible that a difference in sensitivity to different neural processes lead to different response patterns in the two modalities. Although we cannot determine the source of the fMRI/MEG discrepancy from the current data, these explanations would put very strong constraints on the functional interpretation of IFG effects of structure, and therefore the discrepancy itself suggests important and straightforward avenues for future investigation.

## 5 | CONCLUSIONS

A growing body of neurophysiological research has focused on how brain activity in language-sensitive regions of the brain contributes to the processing of sentence-level structure and conceptual-semantic content (Brennan & Pykkänen, 2017; Fedorenko et al., 2016; Halgren et al., 2002; Nelson et al., 2017). In contrast to previous studies, our study identified clear effects of linguistic structure and content at specific sentence positions, rather than averaging across structures or collapsing across syntactic positions and word categories. A major finding of this study is that PTL activity exhibited a robust structure effect that likely reflects predictive syntactic processing: increased activation at the end of the subject noun phrase in sentence context relative to sequences of three unconnected phrases. This effect converges with predictive parsing models of human sentence comprehension (Demberg et al., 2008; Hale, 2001; Lewis & Vasishth, 2005) and provides novel support for the PTL as a critical region for syntax. These relatively coarse observations about the time course of combinatory operations across the sentence provide a potential basis for future integration with more precise neurophysiological models of single word processing derived from ERP responses such as the N400 (Kutas & Hillyard, 1983). Future research should also extend the approach to a broad range of sentence structures in varying contexts in order to determine the extent to which these effects generalize to other syntactic and semantic representations and processing contexts.

## ACKNOWLEDGMENTS

We would like to thank Natalia Lapinskaya & Anna Namyst for their assistance in data collection, and the attendees of the 2015 Society for the Neurobiology of Language Meeting in Chicago and the attendees of the 2018 CUNY Sentence Processing Conference at University of California, Davis for their feedback on this work. The authors have no conflicts of interest.

## ORCID

William Matchin  <https://orcid.org/0000-0002-2530-8429>

## REFERENCES

- Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., & Ochiai, M. (2001). Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Transactions on Applied Superconductivity*, 11(1), 669–672. <https://doi.org/10.1109/77.919433>
- Ahlfors, S. P., Han, J., Belliveau, J. W., & Hamalainen, M. S. (2010). Sensitivity of MEG and EEG to source orientation. *Brain Topography*, 23(3), 227–232.
- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6), 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>
- Bemis, D. K., & Pykkänen, L. (2012). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, 23(8), 1859–1873.
- Bemis, D. K., & Pykkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, 31(8), 2801–2814.
- Binder, J. R. (2012). Task-induced deactivation and the “resting” state. *NeuroImage*, 62(2), 1086–1091.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience*, 11(1), 80–93.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17(1), 353–362.
- Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience*, 15(3), 372–393.
- Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2015). Compositionality and the angular gyrus: A multi-voxel similarity analysis of the semantic composition of nouns and verbs. *Neuropsychologia*, 78, 130–141.
- Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2017). Relational vs. attributive interpretation of nominal compounds differentially engages angular gyrus and anterior temporal lobe. *Brain and Language*, 169, 8–21.
- Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, 26(5), 467.
- Brennan, J., & Pykkänen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *NeuroImage*, 60(2), 1139–1148.
- Brennan, J. R., & Pykkänen, L. (2017). MEG evidence for incremental sentence composition in the anterior temporal lobe. *Cognitive Science*, 41(S6), 1515–1531.
- Brodbeck, C. (2017). Eelbrain: 0.26 [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.322937>
- Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13(6), 407–420. <https://doi.org/10.1038/nrn3241>
- Carroll, L. (1871). *Through the looking glass: And what Alice found there*. London: Macmillan.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29(3), 162–173.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9(2), 179–194.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1), 55–67.
- Demberg, V., Keller, F., & Koller, A. (2013). Incremental, predictive parsing with psycholinguistically motivated tree-adjointing grammar. *Computational Linguistics*, 39(4), 1025–1066.
- Dikker, S., & Pykkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, 127(1), 55–64.
- Fedorenko, E., Hsieh, P. J., Nieto-Castanón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194.
- Fedorenko, E., & Kanwisher, N. (2009). Neuroimaging of language: Why hasn't a clearer picture emerged? *Language and Linguistics Compass*, 3(4), 839–865.
- Fedorenko, E., Nieto-Castanon, A., & Kanwisher, N. (2012). Lexical and syntactic representations in the brain: An fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia*, 50(4), 499–513.
- Fedorenko, E., Scott, T. L., Brunner, P., Coon, W. G., Pritchett, B., Schalk, G., & Kanwisher, N. (2016). Neural correlate of the construction of sentence meaning. *Proceedings of the National Academy of Sciences of the United States of America*, 113(41), E6256–E6262.
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97(20), 11050–11055.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's area. *NeuroImage*, 114, 294–302.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-python. *Frontiers in Neuroscience*, 7, 267.
- Grewe, T., Bornkessel-Schlesewsky, I., Zysset, S., Wiese, R., Von Cramon, D. Y., & Schlesewsky, M. (2007). The role of the posterior superior temporal sulcus in the processing of unmarked transitivity. *NeuroImage*, 35(1), 343–352.
- Hale, J. (2001, June). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies* (pp. 1–8). Association for Computational Linguistics.
- Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage*, 17(3), 1101–1116.
- Hall, E. L., Robson, S. E., Morris, P. G., & Brookes, M. J. (2014). The relationship between MEG and fMRI. *NeuroImage*, 102, 80–91. <https://doi.org/10.1016/j.neuroimage.2013.11.005>
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, 18(4), 665–679.

- Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*, 26(2), 128–138.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York, NY: Oxford University Press.
- Jackendoff, R. (2017). In defense of theory. *Cognitive Science*, 41(S2), 185–212.
- Joshi, A. K., & Schabes, Y. (1997). Tree-adjointing grammars. In *Handbook of formal languages* (pp. 69–123). Berlin, Heidelberg: Springer.
- Kalénine, S., Peyrin, C., Pichat, C., Segebarth, C., Bonthoux, F., & Baci, M. (2009). The sensory-motor specificity of taxonomic and thematic conceptual relations: A behavioral and fMRI study. *NeuroImage*, 44(3), 1152–1162.
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627–633.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11(5), 539–550.
- Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *Journal of Neuroscience*, 33(43), 17174–17181.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.
- Lau, E. F., Weber, K., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2014). Spatiotemporal signatures of lexical-semantic prediction. *Cerebral Cortex*, 26(4), 1377–1387.
- Levy, R. P., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language*, 68(2), 199–222.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375–419.
- Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, 27(1), 1–13.
- Lin, F. H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, 31(1), 160–171.
- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1(4), 476–490.
- Marinkovic, K., Cox, B., Reid, K., & Halgren, E. (2004). Head position in the MEG helmet affects the sensitivity to anterior sources. *Neurology & Clinical Neurophysiology: NCN*, 2004, 30.
- Marks, L. E., & Miller, G. A. (1964). The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 3(1), 1–5.
- Matchin, W. G. (2017). A neuronal retuning hypothesis of sentence-specificity in Broca's area. *Psychonomic Bulletin & Review*, 1–13. <https://doi.org/10.3758/s13423-017-1377-6>
- Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, 88, 106–123.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., ... Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, 5(4), 467–479.
- McKiernan, K. A., D'angelo, B. R., Kaufman, J. N., & Binder, J. R. (2006). Interrupting the "stream of consciousness": An fMRI investigation. *NeuroImage*, 29(4), 1185–1191.
- Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41(5), 329.
- Miller, G. A., & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 2(3), 217–228.
- Münter, T. F., Wieringa, B. M., Weyerts, H., Szentkúti, A., Matzke, M., & Johannes, S. (2001). Differences in brain potentials to open and closed class words: Class and frequency effects. *Neuropsychologia*, 39(1), 91–102.
- Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., ... Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences of the United States of America*, 114(18), E3669–E3678.
- Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, 63(3), 1646–1669.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15(1), 1–25.
- Nobre, A. C., & McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *Journal of Neuroscience*, 15(2), 1090–1098.
- Omaki, A., Lau, E. F., Davidson White, I., Dakan, M. L., Apple, A., & Phillips, C. (2015). Hyper-active gap filling. *Frontiers in Psychology*, 6, 384.
- Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, 108(6), 2522–2527.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago, IL: University of Chicago Press.
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *Journal of Neuroscience*, 35(7), 3276–3284.
- Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, 15(8), 1160–1175.
- Rogalsky, C., Almeida, D., Sprouse, J., & Hickok, G. (2015). Sentence processing selectivity in Broca's area: Evident for structure but not syntactic movement. *Language, Cognition and Neuroscience*, 30(10), 1326–1338.
- Rogalsky, C., & Hickok, G. (2008). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex*, 19(4), 786–796.
- Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2008). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, 19(7), 1493–1503.
- Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either... Or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425.
- Stowe, L. A., Broere, C. A., Paans, A. M., Wijers, A. A., Mulder, G., Vaalburg, W., & Zwartz, F. (1998). Localizing components of a complex task: Sentence processing and working memory. *Neuroreport*, 9(13), 2995–2999.
- Sturt, P., & Lombardo, V. (2005). Processing coordinated structures: Incrementality and connectedness. *Cognitive Science*, 29(2), 291–305.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain* (p. 122). New York, NY: Thieme Medical Publishers.
- Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in Medicine & Biology*, 51(7), 1759–1768.
- Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19(11), 1753–1767.
- Thothathiri, M., Kimberg, D. Y., & Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *Journal of Cognitive Neuroscience*, 24(1), 212–222.
- Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, 14(4), 550–560.
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open and closed-class words. *Memory & Cognition*, 19(1), 95–112.

- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 75(2), 105–143.
- Westerlund, M., Kastner, I., Al Kaabi, M., & Pylkkänen, L. (2015). The LATL as locus of composition: MEG evidence from English and Arabic. *Brain and Language*, 141, 124–134.
- Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia*, 57, 59–70.
- Zaccarella, E., Schell, M., & Friederici, A. D. (2017). Reviewing the functional basis of the syntactic merge mechanism for language: A coordinate-based activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews*, 80, 646–656.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Matchin W, Brodbeck C, Hammerly C, Lau E. The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG. *Hum Brain Mapp*. 2019;40:663–678. <https://doi.org/10.1002/hbm.24403>