

# Simple Composition: A Magnetoencephalography Investigation into the Comprehension of Minimal Linguistic Phrases

Douglas K. Bemis<sup>1</sup> and Liina Pykkänen<sup>1,2</sup>

Departments of <sup>1</sup>Psychology and <sup>2</sup>Linguistics, New York University, New York, New York 10003

The expressive power of language lies in its ability to construct an infinite array of ideas out of a finite set of pieces. Surprisingly, few neurolinguistic investigations probe the basic processes that constitute the foundation of this ability, choosing instead to focus on relatively complex combinatorial operations. Contrastingly, in the present work, we investigate the neural circuits underlying simple linguistic composition, such as required by the minimal phrase “red boat.” Using magnetoencephalography, we examined activity in humans generated at the visual presentation of target nouns, such as “boat,” and varied the combinatorial operations induced by its surrounding context. Nouns in minimal compositional contexts (“red boat”) were compared with those appearing in matched non-compositional contexts, such as after an unpronounceable consonant string (“xkq boat”) or within a list (“cup, boat”). Source analysis did not implicate traditional language areas (inferior frontal gyrus, posterior temporal regions) in such basic composition. Instead, we found increased combinatorial-related activity in the left anterior temporal lobe (LATL) and ventromedial prefrontal cortex (vmPFC). These regions have been linked previously to syntactic (LATL) and semantic (vmPFC) combinatorial processing in more complex linguistic contexts. Thus, we suggest that these regions play a role in basic syntactic and semantic composition, respectively. Importantly, the temporal ordering of the effects, in which LATL activity (~225 ms) precedes vmPFC activity (~400 ms), is consistent with many processing models that posit syntactic composition before semantic composition during the construction of linguistic representations.

## Introduction

Human language derives its unbounded expressive capacity from an ability to take simple, familiar building blocks (such as words) and combine them effortlessly into more complex representations, regardless of whether or not the result is familiar. Such productive “composition” is present in the comprehension of even minimal linguistic expressions. For example, the phrase “blue flamingo” conjures a coherent and meaningful mental representation despite a listener’s presumed unfamiliarity with such an object. Determining the neural bases of basic combinatorial operations, such as those underlying this simple construction, is therefore a central goal for the cognitive neuroscience of language and is a necessary precursor to understanding the comprehension of more complex expressions.

Interestingly, rather than focus on such elementary combinatorial operations, historically neurolinguistic studies have been driven primarily by questions about how people handle deviations from simple processing. Past experiments typically investigate complex syntactic structures, such as center-embedded or

object-extracted relative clauses (Miller and Chomsky, 1963; Stromswold et al., 1996), or unexpected phrases, such as garden-path constructions (Bever, 1970; Osterhout et al., 1994) and implausible sentence completions (Kutas and Hillyard, 1980). In these manipulations, basic linguistic expressions serve as the baseline, and effects are determined as differences in neural activity engendered by the more complex or surprising stimuli. Consequently, in these studies, information about basic combinatorial operations can only be gained indirectly by arguing that effects observed during the comprehension of more complex stimuli reflect differential processing by the same mechanisms responsible for comprehending basic constructions, a presupposition that may or may not be correct.

The few neurolinguistic investigations that reverse this norm and place basic combinatorial operations at the forefront have done so within the context of full sentences. Most often these studies compare the comprehension of sentences with that of unstructured lists of words or pseudowords (Mazoyer et al., 1993; Friederici et al., 2000). Although effects elicited through such comparisons likely reflect basic combinatorial processes, it is unclear to what extent such results also include additional comprehension mechanisms, such as the establishment of reference and long-distance dependencies, the temporal anchoring of events, pragmatic inferences, the deployment of various types of memory, and potentially the construction of discourse representations (because most studies perform their analysis over multiple sentences at a time).

In the present work, we reduce the complexity of the critical stimuli to simple adjective–noun phrases (e.g., “red boat”) to

Received Sept. 23, 2010; revised Nov. 16, 2010; accepted Dec. 1, 2010.

This research was supported by the National Science Foundation Grant BCS-0545186 (L.P.), the Neuroscience of Language Laboratory at New York University (NYU) Abu Dhabi award from the NYU Abu Dhabi Research Council (L.P.), and the Whitehead Fellowship for Junior Faculty Biomedical and Biological Sciences (L.P.). We thank Anna Bemis for her assistance in creating the figures.

Correspondence should be addressed to Douglas K. Bemis, New York University, 6 Washington Place, New York, NY 10003. E-mail: doug.bemis@nyu.edu.

DOI:10.1523/JNEUROSCI.5003-10.2011

Copyright © 2011 the authors 0270-6474/11/312801-14\$15.00/0

investigate the neurophysiological foundations of elementary combinatoric processing. Using magnetoencephalography (MEG), we assessed increases in neural activity generated within minimal compositional contexts compared with matched non-compositional environments (see Fig. 1). The temporal resolution of MEG allowed us to examine combinatorial processing in specific, well-controlled linguistic contexts in which only the presence or absence of composition was varied. Provocatively, areas canonically implicated in traditional neurophysiological models of language processing appear to play a lesser role in basic composition. In contrast, two areas recently discussed in numerous neurobiological studies, anterior temporal cortex and ventromedial prefrontal cortex (vmPFC), are implicated in this critical and fundamental construction of linguistic structure and meaning.

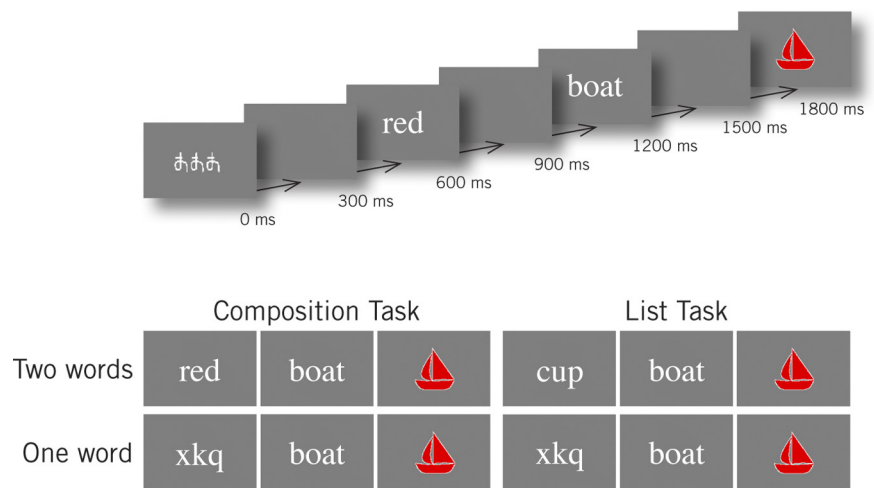
## Materials and Methods

**Participants.** Twenty-five non-colorblind, native English speakers participated in the study (17 female; average age, 25.3 years). All had normal or corrected-to-normal vision and gave informed consent.

**Experimental design and stimuli.** Our paradigm was designed to isolate neural activity associated with basic combinatorial mechanisms that operate during the comprehension of very simple linguistic phrases. We measured activity generated by the processing of common, object-denoting nouns presented in either a combinatorial phrasal context or a non-combinatorial control context. Lexical operations elicited by the comprehension of multiple linguistic items (as in a compositional phrase) were controlled for by having participants perform two tasks, each consisting of a two-word and a one-word condition. The presence or absence of composition was then varied between tasks (Fig. 1).

In the main “composition” task, combinatorial contexts consisted of a color adjective followed by a simple noun (“red boat”), whereas in non-combinatorial contexts a length-matched, unpronounceable consonant string preceded the noun (“xkq boat”). To ensure attention to the critical words, participants were asked to indicate whether a following colored shape matched the preceding verbal material. To count as matching in the two-word trials (“red boat”), both the color and the shape of the picture had to match the preceding words. In the one-word trials (xkq boat) only the shape of the picture was relevant for the decision.

Because almost any meaningful word before the noun might elicit attempts at composition in this task, to control for the difference in lexical–semantic material preceding the noun in these two conditions, we had participants perform a separate “list” task, also consisting of two-word and one-word trials. In this control task, participants saw either two object-denoting nouns (“cup, boat”) or an unpronounceable consonant string followed by a noun (“xkq boat”), as before. Importantly, participants were now instructed to decide whether the subsequent colored shape matched any of the words that preceded it rather than all of the words, as before. Because the critical stimuli (e.g., “boat”) were held constant across all four conditions, this control task allowed us to assess whether increases in activity observed during the two-word composition condition compared with the one-word composition condition might simply be attributable to the presentation of two words as opposed to one, because similar differences would be expected to occur within this control task as well. Note that, although noun–noun compounds are grammatical phrases in English, the list task discouraged participants from processing the two-word trials as noun–noun compounds, because they were required to assess the following shape against either one noun or the other. Thus, our design crossed task (composition vs list) and number of words (one vs two), with the expectation that only the two-



**Figure 1.** Experimental design. Our design crossed task (composition vs list) and number of words (two vs one). In each trial, participants indicated whether the target picture matched the preceding words. To satisfy this criterion, in the composition task, all preceding words were required to match, whereas in the list task, any matching word sufficed. A total of six colors and 25 shapes were randomly combined and used as stimuli. Half of the target pictures matched, but half did not. Only activity recorded at the matched nouns (“boat”) was analyzed.

word composition trials should elicit activity related to linguistic combination.

Each trial contained a fixation, an initial word or non-word, a critical noun, and a target shape. The initial word or non-word varied by condition and could be an adjective (two-word composition condition), noun (two-word list condition), or unpronounceable consonant string (one-word conditions). Twenty-five one-syllable nouns were used in the second (critical) stimulus position (disc, plane, bag, lock, cane, hand, key, shoe, bone, square, bell, boat, bow, car, cross, cup, flag, fork, heart, lamp, leaf, note, star, tree, house). Adjectives in the two-word composition condition (red, blue, pink, black, green, brown) were matched in length with the first stimuli nouns used in the list task (cup, boat, lamp, plane, cross, house). In both tasks, one-word trials were produced by substituting each adjective or initial noun with an unpronounceable consonant string of the same length (xkq, qxsw, mtpv, rjdnw, wvcnz, zbxlv). The sets of adjectives, non-words, and first-stimuli list nouns were each created in 28-point non-proportional Courier font and were all matched in total number of pixels (mean  $\pm$  SD, 420.17  $\pm$  77.39, 415.12  $\pm$  49.40, and 420.33  $\pm$  76.74, respectively). Target shapes were manually created to depict a canonical, unambiguous representation of one of the nouns, filled in with one of the six colors denoted by the adjectives. Three versions of each target were then created by applying a random scaling factor between 105 and 115% and a random rotation of 0–360° to the original figure. All stimuli were presented using PsyScope 1.2.5 (Cohen et al., 1993) and were projected  $\sim$ 50 cm from the participant’s eye. Words subtended between 2° and 4°, whereas target shapes were larger, subtending between 6° and 10°.

Both tasks also used a second, one-word control condition in addition to the unpronounceable consonant strings. This condition used novel symbol strings as the first stimuli (in place of the consonant strings) and was intended to minimize, to the greatest extent possible, the chance that participants attempted composition between the initial stimulus and the following noun. Thus, for these stimuli, we created repeated symbol strings that carried no obvious linguistic association and were length-matched to the adjectives (ㄨㄨㄨ, 4444, AAAAA, ㄸㄸㄸㄸ, XXXXXX, ㄩㄩㄩㄩ). We then repeated these stimuli as the fixations for all conditions to prevent any surprisal or interest-related activity when they were substituted for the adjectives. As with the non-words and initial list task nouns, these symbol strings were matched in number of discrete objects, filled area, and pixel density to the adjectives used in the two-word composition condition. However, because the results for this condition were qualitatively identical to the one-word condition described in the text, for simplicity, discussion of this third condition has been eliminated from the main body of the paper.

During each task, participants viewed 300 trials, 100 of each trial type. All conditions contained an equal number of trials in which the target shape matched or did not match the preceding words, whereas the two-word composition condition additionally divided the non-matching trials equally among those that did not match the adjective and those that did not match the noun. In these trials, all target shapes matched at least one of the preceding words. During each condition, each of the 25 critical nouns was used four times, twice in matching and twice in non-matching trials. The same set of colored shapes was used as targets for each set of matching and non-matching trials in each condition in each task, such that only 50 different colored shapes were used in total during each experimental run. However, the trials and stimulus lists were randomized and constructed separately for each participant.

**Procedure.** Before the experiment, participants practiced their first task outside of the MEG room. Although participants were made aware of the existence of a second task at this time, no specific instructions regarding the second task were given before the completion of the first task. Instructions and practice for this second task were then given after the completion of the first task, while participants were in the machine. Before recording, participants' head shapes were digitized using a Polhemus Fastrak three-dimensional digitizer. The digitized head shape was then used to constrain source localization during analysis by coregistering five coils located around the face with respect to the MEG sensors.

During the experiment, participants lay in a dimly lit, magnetically shielded room and performed the two tasks in separate blocks. The order of tasks was counterbalanced between participants. MEG data were collected using a whole-head 157-channel axial gradiometer system (Kanazawa Institute of Technology, Nonoichi, Japan) sampling at 1000 Hz with a low-pass filter at 200 Hz and a notch filter at 60 Hz. During each trial, all stimuli besides the target shapes were presented for 300 ms and were followed by a 300 ms blank screen. Target shapes appeared at the end of each trial and remained onscreen until the participant made a decision. Subsequent trials began after a blank screen was shown for a variable amount of time. This delay followed a normal distribution with a mean  $\pm$  SD of  $400 \pm 100$  ms. The recording lasted  $\sim 45$  min.

**Data acquisition.** MEG data from the 100 ms before the onset of each critical noun to 500 ms after onset were segmented out for each participant for each condition. Raw data were first cleaned of potential artifacts by rejecting trials for which the participant answered either incorrectly or too slowly (defined as over 2.5 s after the appearance of the target shape) or for which the maximum amplitude exceeded a threshold that varied between 2500 and 3500 fT depending on the amplitude range of each participant. Overall,  $8.8 \pm 5.6$  (mean  $\pm$  SD) of trials were excluded. Remaining data were then averaged for each participant for each condition and bandpass filtered between 1 and 40 Hz. For inclusion in additional analysis, we required that participants show a qualitatively canonical profile of evoked responses during the processing of the critical items. This profile was defined as the appearance of robust and prominent initial visual responses, either the M100 or M170 field pattern (Tarkiainen et al., 1999; Pylkkänen and Marantz, 2003), in the time window of 100–200 ms after the critical stimuli. To assess this criterion, preliminary grand average waveforms were constructed for each participant by averaging over all conditions. Five participants failed to meet this requirement and were excluded from additional analysis.

**Minimum norm estimates.** Distributed minimum norm source estimates served as our primary dependent measure. After preprocessing, a source estimate was constructed for each condition average using L2 minimum norm estimates, calculated in BESA 5.1 (MEGIS Software GmbH). The channel noise covariance matrix for each estimate was based on the 100 ms before the onset of the noun in each condition average. Each minimum norm estimate was based on the activity of 1426 regional sources evenly distributed in two shells 10 and 30% below a smoothed standard brain surface. Regional sources in MEG can be regarded as sources with two single dipoles at the same location but with orthogonal orientations. The total activity of each regional source was then computed as the root mean square (RMS) of the source activities of its two components. Pairs of dipoles at each location were first averaged, and then the larger value from each source pair was chosen, creating 713 non-directional sources for which activation could be compared across

participants and conditions. Minimum norm images were depth weighted as well as spatiotemporally weighted, using a signal subspace correlation measure (Mosher and Leahy, 1998).

**Data analysis.** To determine effects related to basic combinatorial processing, our primary analysis examined source activity localized to specific regions of interest (ROIs) selected from either studies similar in focus or technique to the present investigation or historically significant neurophysiological models. To verify the extent and robustness of these results, we also performed a full-brain comparison across all source and time points followed by an analysis of the sensor data targeted at assessing the reproducibility in sensor space of the significant effects identified by the distributed source analyses.

**Regions of interest.** Regions of interest were drawn on the smooth cortex used to reconstruct estimated cortical activity from the MEG data (see Fig. 4). ROIs derived from past functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) paradigms were given relatively broad interpretations, to compensate for the reduced spatial resolution of MEG.

**Anterior temporal lobe ROIs.** The studies most similar in spirit to the present investigation are those that contrast the processing of sentences to that of non-combinatorial stimuli (e.g., unstructured word lists or environmental sounds). Compared with these baselines, the processing of sentences has consistently shown increased activity in the left anterior temporal lobe (LATL) (Mazoyer et al., 1993; Bottini et al., 1994; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2001, 2005, 2006; Vandenberghe et al., 2002). The majority of these studies have also identified increased activity within the right anterior temporal lobe (RATL), although nearly always to a lesser degree (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2001). Effects in both hemispheres primarily localize to the anterior temporal pole [approximately Brodmann's area (BA) 38], although many studies have also found increased activity in adjacent regions of the anterior middle and anterior inferior temporal gyri (Mazoyer et al., 1993; Bottini et al., 1994; Stowe et al., 1998; Humphries et al., 2006). Thus, to assess the role of the anterior temporal lobes in basic combinatorial processing, we placed an ROI in both the LATL and RATL, each encompassing approximately the area around BA 38 and the anterior portions of BA 20 and 21 (see Fig. 4).

**Ventromedial prefrontal cortex ROI.** Within MEG, a series of studies has investigated the neural correlates of semantic composition through manipulations involving semantic mismatches (Pylkkänen and McElree, 2007; Brennan and Pylkkänen, 2008, 2010; Pylkkänen et al., 2009b). In these studies, expressions are contrasted for which the computations necessary to construct a coherent meaning are varied while the amount of syntactic work is maximally controlled. The processing of more semantically involved mismatched expressions has consistently elicited increased activity in an MEG component, named the anterior midline field (AMF), that peaks at  $\sim 400$  ms and localizes to the vmPFC. Activity in the vmPFC also increases when participants attempt to resolve semantic violations but not when they encounter similar violations of world knowledge (Pylkkänen et al., 2009a). The exact location of the vmPFC effect has been somewhat variable across these studies, perhaps because of the difficulty inherent in localizing deep cortical sources in MEG. The probability of correct localization decreases with the depth of the source in MEG (Hillebrand and Barnes, 2002). Thus, there may be both increased variability and decreased power when measuring activity from a source that is more distant from the sensors, such as the vmPFC. However, the majority of previous effects have localized to the medial region of the orbitofrontal cortex (Pylkkänen and McElree, 2007; Pylkkänen et al., 2009a,b; Brennan and Pylkkänen, 2010), with activity occasionally spreading either more ventrally (Brennan and Pylkkänen, 2010) or dorsally (Brennan and Pylkkänen, 2008; Pylkkänen et al., 2009b). Thus, to assess the contribution of the vmPFC to basic linguistic composition, we placed an ROI within the whole of the medial, ventral portion of the prefrontal cortex (i.e., approximately the entirety of BA 11) (see Fig. 4).

**Traditional neurolinguistic ROIs.** Historically, by far the most extensively investigated regions in neurolinguistics have been, approximately, the left inferior frontal gyrus (LIFG) ("Broca's area") and a broad region around the left posterior middle and superior temporal gyri [i.e., the left posterior temporal lobe (LPTL); "Wernicke's area"]. Together, these two

areas constitute the “traditional” neurophysiological language network. The history of these regions extends well into the 19th century (Broca, 1861; Wernicke, 1874) and has yet to coalesce into any resemblance of a consensus regarding their function or even precise location. Despite the paucity of strong evidence directly tying either region to basic linguistic combinatorial processing, we included these familiar regions in our analysis because of their ubiquitous presence throughout neurophysiological studies of language.

The LIFG has frequently been implicated in the processing of many types of complex syntactic constructions, such as center embedding (Stromswold et al., 1996; Caplan et al., 2000), wh-extraction (Santi and Grodzinsky, 2007), and scrambled verbal arguments (Friederici et al., 2006). Along with these syntactic findings, the LIFG has also been implicated in many processes that are not strictly syntactic in nature, for example, the maintenance and manipulation of working memory (Cohen et al., 1997; Fiebach et al., 2007), cognitive control (Badre and Wagner, 2007), and executive functioning (Koechlin and Summerfield, 2007). Because of the breadth of this empirical coverage, the LIFG has been suggested to subserve any number of mental mechanisms from the narrowly specific [e.g., syntactic movement only (Grodzinsky and Santi, 2008)] to the broadly general [e.g., selection and regulation among competing mental representations (Thompson-Schill et al., 2005)]. Although the exact placement of the LIFG has varied alongside its multifaceted interpretations, in general, the pars opercularis and pars triangularis (approximately BA 44 and BA 45) have served as the center for these investigations. Therefore, we created a relatively broad ROI around these regions from which to measure LIFG activity in the present study (see Fig. 4).

The LPTL region has become decidedly more fragmented over the years, dividing into several apparently functionally distinct areas. The posterior portion of the middle temporal gyrus (pMTG) has been most heavily implicated in lexical-level operations (i.e., the storage and retrieval of long-term linguistic information). Activity measured by fMRI within the pMTG exhibits lexical priming effects across a wide range of stimulus-onset asynchronies and modalities (Gold et al., 2006) and has been found to increase with the number of words processed per trial (Badre et al., 2005) and for judgments requiring lexical–semantic information (Gitelman et al., 2005). Within MEG, activity localized to this general region has shown sensitivity to many factors that affect the timing of lexical access, such as frequency (Embick et al., 2001), repetition (Pylkkänen et al., 2000), and phonotactic probability (Pylkkänen et al., 2002) (for review, see Pylkkänen and Marantz, 2003; Salmelin, 2007). Thus, there is converging evidence from both the hemodynamic and electrophysiological literature that the pMTG plays an integral role in the storage and retrieval of lexical information. In relation to combinatorial operations, it has been suggested that classic N400 effects, canonically elicited by semantically incongruous sentence completions, e.g., “He spread the warm bread with socks” (Kutas and Hillyard, 1980; Hagoort et al., 2004), localize, at least partially, to the pMTG (for review, see Lau et al., 2008). However, the exact relationship between these N400 effects and combinatorial processing is currently unclear, because the N400 is also modulated by many lexical-level factors, such as frequency (Van Petten and Kutas, 1990; Allen et al., 2003) and repetition (Rugg, 1985), and can be elicited within non-compositional contexts, such as semantically unrelated word pairs (Holcomb, 1993).

Similar ambiguity surrounds an adjacent region of the LPTL, the angular gyrus (AG), located superior and posterior to the pMTG (approximately BA 39). The AG has been suggested recently to play a role in the integration of semantic information into context (Lau et al., 2008), because increased AG activity has been observed in response to semantically congruous sentences compared with semantically related word lists or syntactically well-formed but meaningless sentences (Humphries et al., 2006, 2007). Also, the AG has occasionally exhibited more activity during the reading of sentences compared with word lists (Bottini et al., 1994; Bavelier et al., 1997). However, the AG has been implicated in lexical-level processing as well, showing increased activity during the presentation of single words compared with non-words during both lexical decision tasks (Rissman et al., 2003; Ischebeck et al., 2004) and simple reading (Bookheimer et al., 1995; Binder et al., 2005).

Despite the fractionation and uncertainty surrounding the LPTL, for the sake of simplicity, we enclosed both the pMTG and the AG within a single LPTL ROI to identify any composition-related activity in this region (see Fig. 4). (However, it should be briefly noted here that dividing this region into two distinct ROIs does not substantively affect the results.)

**ROI analysis.** Significant effects within each ROI were determined by applying a nonparametric, cluster-based permutation test (Maris and Oostenveld, 2007) to the entire time interval surrounding the onset of the critical noun (0–500 ms). For this analysis, we constructed a test statistic capable of isolating clusters of time points for which the activity localized to each ROI exhibited the pattern we hypothesized for composition-related activity. Specifically, we identified periods of contiguous time points for which the increase in activity during the two-word composition condition compared with the one-word composition condition was significantly greater than any activity difference seen between the two list task conditions. ROIs exhibiting such effects were then further investigated using a follow-up permutation test within each task. This follow-up test used a similar but simpler test statistic applicable for identifying increases in one condition relative to another.

The logic for the permutation test is as follows. First, a cluster-based test statistic for the observed data is calculated (as described below). Then, the same test statistic is calculated for many permutations of the actual data, creating a distribution for the statistic. Permuted datasets are created by randomly assigning the condition labels within each participant. Importantly, this assignment is done independently for each participant and only once per participant per permutation (as opposed to once for each time point for each participant). The *p* value of the observed test statistic is then computed relative to this distribution and is set equal to the proportion of permuted datasets that produce a test statistic more extreme than that of the actual data.

Our experimental design was constructed such that composition-related activity should exhibit a particular type of interaction between task and number of words. Specifically, within the composition task, we expected composition-related activity to be greater within the two-word condition compared with the one-word condition, whereas within the list task, we expected to see no composition-related activity in either condition. Thus, following the discussion by Maris and Oostenveld (2007, p 187), we used a test statistic for our permutation test that incorporates this previous hypothesis and is therefore capable of identifying these particular types of interactions. First, we identified clusters of adjacent time points for which there was an interaction between task and number of words by submitting the data at each time point to a  $2 \times 2$  repeated-measures ANOVA with task (composition vs list) and number of words (one vs two) as factors. Clusters were then extracted by identifying intervals of at least 10 adjacent time points for which, at each time point, the resulting *p* value for the interaction between the two factors was less than a set threshold, chosen so as to identify relatively shallow yet long-lasting patterns of activity ( $p = 0.30$ ). Then, at each time point in a given cluster, we performed a paired *t* test within each task. This gave us two *t* values for every time point in the cluster: one corresponding to the difference in activity between the two composition conditions and one corresponding to the difference in activity between the list conditions. The test statistic for a cluster was then calculated by summing all composition *t* values and subtracting the magnitude of all list *t* values. The use of this test statistic allowed us to specifically identify clusters for which the interaction observed in the initial phase of the test was driven by two concurrent factors: (1) an increase in activity during the two-word composition condition relative to the one-word composition condition and (2) no difference in activity between the list task conditions. In other words, this test statistic, and consequently the test itself, incorporates our previous hypothesis regarding the particular type of interaction in which we are interested and so can uncover effects directly relevant to this hypothesis.

For follow-up permutation tests within each task, clusters were identified by performing a paired *t* test between the one-word and two-word activity measures at each time point and then extracting intervals for which the resulting *p* value was less than the same set threshold as before ( $p = 0.30$ ) for a minimum of 10 adjacent time points. The test statistic

for each cluster was then calculated simply as the sum of each  $t$  value for each time point in the cluster. For all tests, the reported  $p$  value was obtained from a distribution of test statistics derived from 10,000 permutations of the original data.

**Whole-brain analysis.** Whole-brain analyses were performed within each task. Two-word and one-word activity measures were compared sample by sample for every source time point using a paired  $t$  test. A difference was considered significant if it remained reliable ( $p < 0.05$ ) for at least 10 samples (10 ms) and was observed in at least 10 adjacent cortical sources. In the results and figures below, we discuss only effects attributable to an increase in two-word activity compared with one-word activity within each task.

**Sensor space analysis.** Our sensor space analysis was designed to determine the extent to which effects identified through the source space analyses were also apparent in the sensor data. Thus, we performed one sensor space analysis for each significant, combinatorial source space effect as determined by the ROI analyses. For each such effect, we first identified a set of sensors of interest for each participant based on the spatiotemporal location of the source space result. Then, we performed the same cluster analysis as for the source space ROIs on the RMS of the data from these sensors.

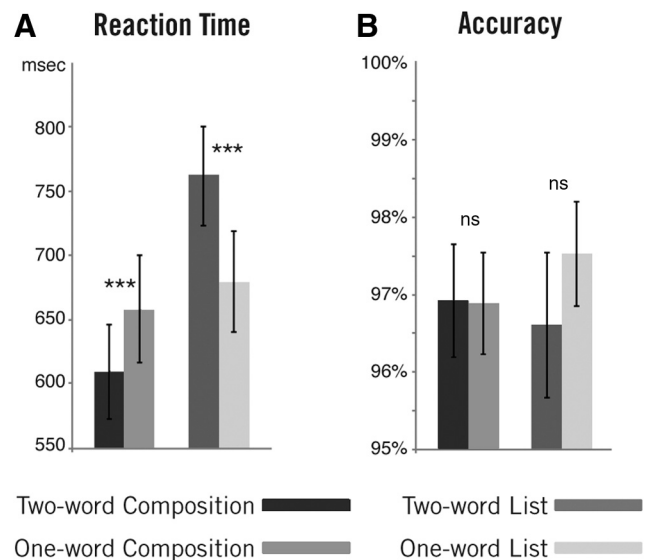
Because the mapping between source space and sensor space was not clear a priori, we adopted a split-half, cross-validation method to analyze the sensor data. Initially, the data were divided into two halves by randomly partitioning the trials used for source reconstruction, for each participant for each condition. We then reaveraged each data partition, producing two equally sized sets of data, each containing an average per participant per condition. As during source reconstruction, these averages were filtered between 1 and 40 Hz for analysis. Sensors of interest for each participant were then selected based on the data from the two-word composition condition. In one-half of the data, the training set, we first visually inspected the grand average sensor data for the two-word composition condition, accumulated from all of the participants, and identified the dominant field pattern present in this data during the peak of the source space effect. Then, for each participant, we inspected their individual two-word composition condition average, also from the training set, and identified the set of sensors whose data most closely resembled this dominant component. In each case, we chose the sensors in the centers of the ingoing and outgoing fields of the dominant component and the set of sensors surrounding these peaks for which the field pattern remained strongly present. Across all selections, this resulted in an average  $\pm$  SD of  $12.95 \pm 4.065$  sensors per set (minimum, 6; maximum, 25). If the dominant component was not apparent in the participant's average, that participant was excluded from the current analysis. To create the dependent measure for input into the cluster tests, we then calculated the RMS for the set of sensors chosen for each participant using the other half of the data, the testing set. Thus, the data used to select the sensors was independent of the data used in the analysis. We applied the cluster test to this data over the entire interval (0–500 ms) as before and also to a more circumscribed time interval surrounding the extent of the corresponding source space effect, to compensate for the loss of power that results from partitioning the data. Finally, we switched the training and test sets and repeated the analysis for each source space effect. Thus, ultimately, we produced two complementary sensor space analyses for each source space effect of interest.

## Results

### Behavioral results

Accuracy and reaction time data for responses were submitted to a  $2 \times 2$  repeated-measures ANOVA with task (composition vs list) and number of words (one vs two) as factors (Fig. 2). There were no significant effects found for accuracy, for task, number of words, or their interaction (all  $F$  values  $< 1$ ), because participants were generally close to ceiling on all conditions (average  $\pm$  SD: two-word composition,  $96.9 \pm 3.2\%$ ; one-word composition,  $96.9 \pm 2.9\%$ ; two-word list,  $96.6 \pm 4.2\%$ ; one-word list,  $97.5 \pm 3.0\%$ ).

We found a strong interaction between task and number of words for reaction times, ( $F_{(1,19)} = 89.2$ ,  $p < 0.001$ ). Paired  $t$  tests

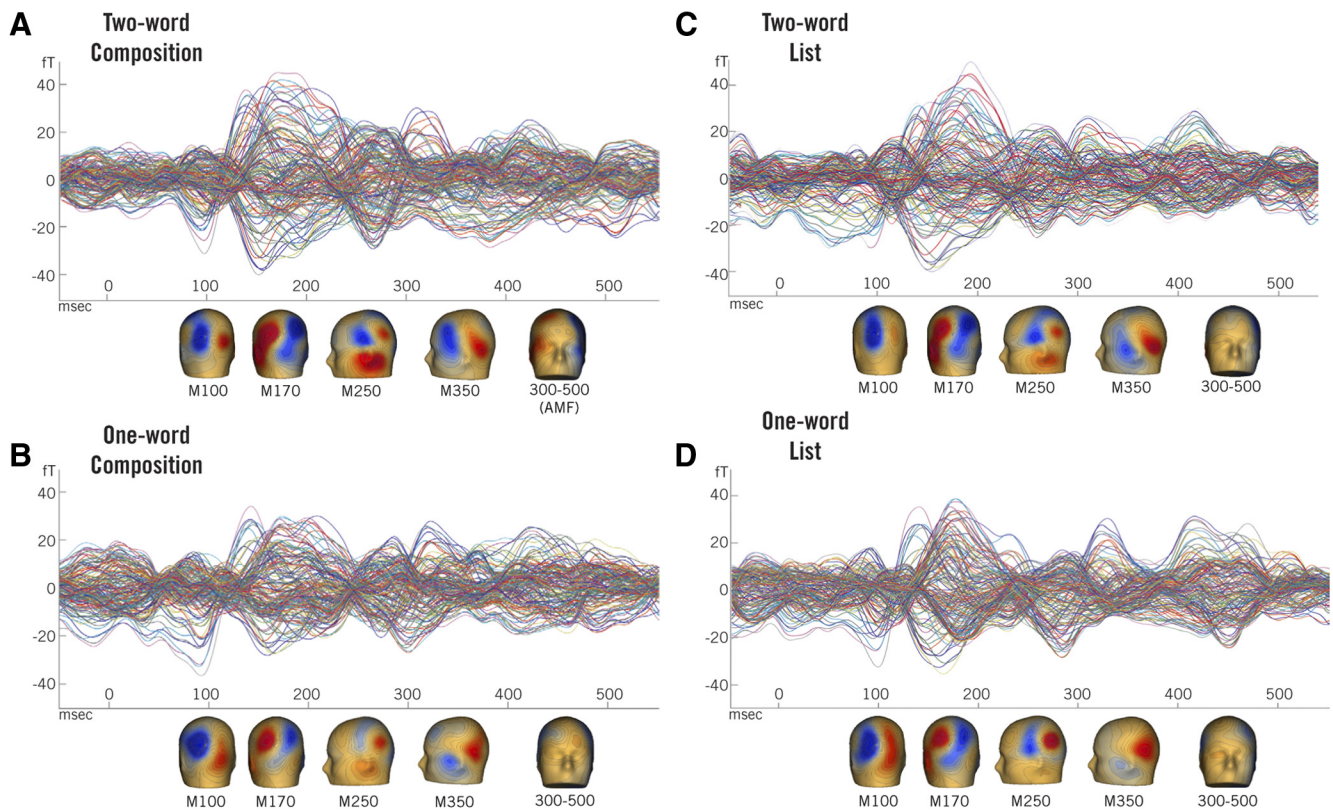


**Figure 2.** Behavioral results. Reaction time (**A**) and accuracy (**B**) data were submitted to a  $2 \times 2$  repeated-measures ANOVA with task (composition vs list) and number of words (one vs two) as factors. We observed a significant interaction between the two factors for reaction time ( $F_{(1,19)} = 89.2$ ), with *post hoc* tests revealing slower responses in the two-word list condition and faster responses in the two-word composition condition compared with matched one-word controls. No significant effects were found for accuracy. ns, Nonsignificant; \*\*\* $p < 0.001$ .

within each task revealed that, compared with the corresponding one-word conditions, participants were markedly slower for the two-word condition within the list task (average  $\pm$  SD,  $762 \pm 173$  vs  $679 \pm 173$  ms;  $p < 0.001$ ), whereas they were significantly faster for the two-word condition in the composition task (average  $\pm$  SD,  $609 \pm 165$  vs  $658 \pm 186$  ms;  $p < 0.001$ ). These results indicate that, in the list task, the two-word condition was more difficult than the one-word condition, because participants were much slower to respond and were slightly less accurate in the former compared with the latter. Because participants needed to remember twice as many items in the two-word condition compared with the one-word condition, this result is relatively unsurprising and echoes many past studies that demonstrate that response times in memory tasks increase with the number of items to remember (Sternberg, 1967). Conversely, in the composition task, the two-word condition appeared to be the easier of the two, because participants responded significantly faster in this condition than in the one-word condition. This result is somewhat surprising because participants had to assess the correspondence of twice as many features in the two-word condition compared with the one-word condition at the presentation of the target shape. However, this effect is in line with previous results demonstrating that, after an adjective–noun description, participants could identify pictures that incorporated the adjective faster than those that did not (Potter and Faulconer, 1979). Importantly, the present behavioral results suggest that any increase in neural activity observed for the two-word composition condition does not reflect increased effort because this critical condition appeared to be the easiest of the four.

### General assessment of MEG sensor data

A qualitative overview of the evoked MEG response during the presentation of the critical nouns is shown in Figure 3. The initial baseline activity preceding the onset of the noun (used to estimate the noise covariance matrix during the construction of each minimum norm estimate) showed no significant differences in



**Figure 3.** MEG sensor data. The average evoked response to the critical noun is displayed for each condition: **A**, two-word composition; **B**, one-word composition; **C**, two-word list; **D**, one-word list. Canonical visual response peaks and field patterns (M100 and M170) (Tarkiainen et al., 1999; Pykkänen and Marantz, 2003) are visible at ~100 and 150 ms in all four conditions. Subsequent M250 and M350 peaks and field patterns (Pykkänen et al., 2002; Pykkänen and Marantz, 2003) are also present across all four conditions. A typical anterior midline field pattern (Pykkänen and McElree, 2007) can be seen accompanying a sustained increase in activity from 300 to 500 ms within the two-word composition condition. This field pattern is absent from the other three conditions.

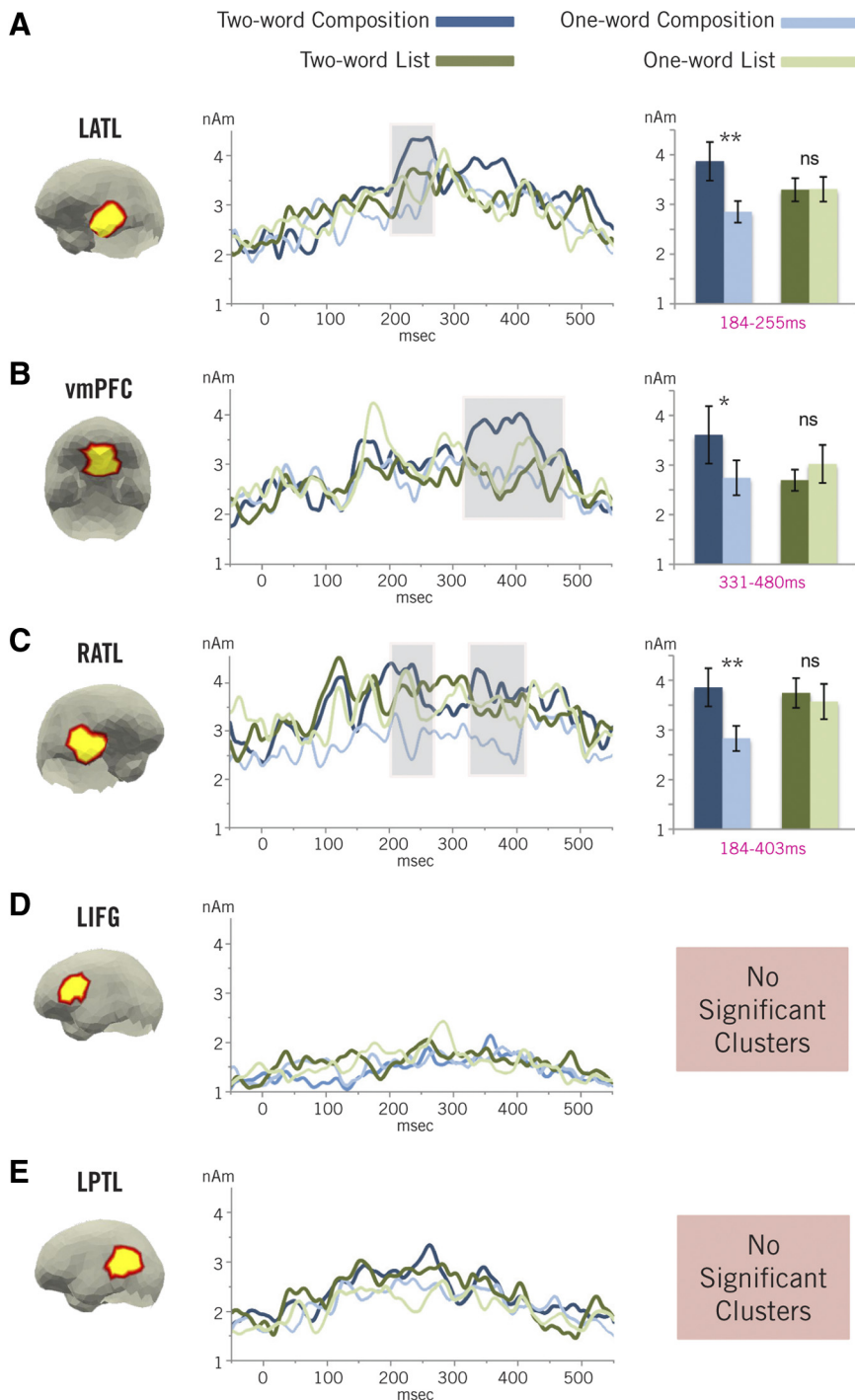
average amplitude between conditions (task  $\times$  words two-way repeated-measures ANOVA yielded all  $F$  values  $< 0.5$ ; average  $\pm$  SD: two-word composition,  $16.09 \pm 3.89$  fT; one-word composition,  $16.60 \pm 4.86$  fT; two-word list,  $16.65 \pm 5.72$  fT; one-word list,  $16.62 \pm 4.52$  fT). Across all conditions, we observed canonical early visual responses at ~100 and 150 ms that have consistently been identified in MEG during the presentation of visual words (Tarkiainen et al., 1999; Pykkänen and Marantz, 2003). These visual responses were followed by the characteristic M250 and M350 field patterns, focused over the left temporal lobe, that also have been consistently observed after visually presented words (Embick et al., 2001; Pykkänen et al., 2002; Pykkänen and Marantz, 2003). Additionally, a large, sustained increase in MEG activity can be seen in the two-word composition condition beginning at ~350 ms and continuing to 450 ms. This activity is accompanied by the characteristic AMF field pattern (Pykkänen and McElree, 2007), which we did not observe in any of the control conditions during this time.

#### Anterior temporal lobe ROIs

Significant clusters of time points were identified by the interaction permutation test within both the LATL and RATL (Fig. 4). Within the LATL, a significant cluster of combinatorial activity was found from 184 to 255 ms ( $p = 0.0389$ ; average  $\pm$  SD: two-word composition,  $3.87 \pm 1.74$  nanoampere meters (nAm) one-word composition,  $2.85 \pm 0.97$  nAm; two-word list,  $3.29 \pm 1.04$  nAm; one-word list,  $3.31 \pm 1.11$  nAm). A follow-up test within the composition task alone revealed a temporally similar cluster of time points for which activity increased during the two-word condition compared with the one-word condition, al-

though this cluster was only marginally significant (204–263 ms;  $p = 0.071$ ; two-word,  $4.13 \pm 1.97$  nAm; one-word,  $3.03 \pm 1.20$  nAm). Within the list task, no clusters of increased activity during the two-word condition were seen at any time within the LATL (all clusters  $p > 0.35$ ). Inspecting the waveforms from this ROI, there appears to be an additional increase in two-word composition activity relative to the other three conditions, peaking at 375 ms (Fig. 4). However, despite the visual clarity of this separation, activity during this time did not approach significance in the interaction cluster test (closest cluster, 327–334;  $p = 0.5549$ ). A follow-up test within the composition task alone also failed to approach significance (closest cluster, 360–380;  $p = 0.4218$ ). Even a direct comparison of two-word to one-word composition activity levels from 350 to 400 ms using a one-tailed paired  $t$  test only approached significance ( $p = 0.076$ ; average  $\pm$  SD: two-word,  $3.75 \pm 1.73$  nAm; one-word,  $3.13 \pm 1.56$  nAm). Thus, despite the tantalizing nature of this visual separation, no firm conclusions can be drawn at this time about activity within the LATL during this later time window. There is, however, strong evidence that activity in the LATL reflects basic combinatorial processing in the earlier time window, from 184 to 255 ms.

Within the RATL, two highly significant clusters were identified by the initial permutation test: 184–246 ms ( $p = 0.0078$ ; average  $\pm$  SD: two-word composition,  $3.86 \pm 1.72$  nAm; one-word composition,  $2.83 \pm 1.13$  nAm; two-word list,  $3.73 \pm 1.34$  nAm; one-word list,  $3.57 \pm 1.58$  nAm) and 329–403 ms ( $p = 0.0041$ ; average  $\pm$  SD: two-word composition,  $4.26 \pm 2.02$  nAm; one-word composition,  $2.86 \pm 1.12$  nAm; two-word list,  $3.69 \pm 1.53$  nAm; one-word list,  $3.67 \pm 1.93$  nAm). A follow-up test

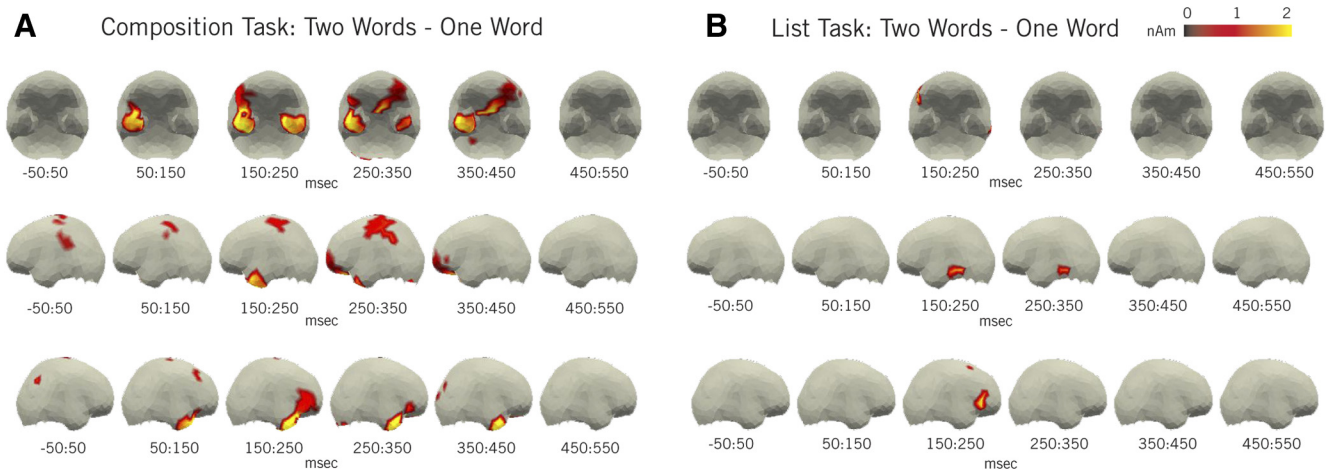


**Figure 4.** ROI results. Localized activity is shown for the five ROIs during the comprehension of the critical nouns, averaged across participants. Shaded regions denote significant clusters of combinatorial activity as identified by a cluster-based permutation test (Maris and Oostenveld, 2007) applied to the entire interval of 0–500 ms. Within the LATL (A), a significant cluster of combinatorial activity was found from 184 to 255 ms ( $p = 0.039$ ), within which two-word composition activity was significantly greater than one-word composition activity, and activity in the list conditions did not differ. A later cluster of combinatorial activity was found in the vmPFC (B) from 331 to 480 ms ( $p = 0.014$ ). In this time period, localized activity was again significantly greater for the two-word composition condition compared with the one-word composition condition. However, unlike in the LATL and vmPFC, activity in this region also showed a significant difference between the two one-word conditions during the identified time window ( $p = 0.041$ ), suggesting a task-related suppression of activity in the one-word composition condition as opposed to an increase in activity during the two-word composition condition. No significant clusters of combinatorial activity were found in either the LIFG (D) or LPTL (E) ROIs. Targeted tests within the composition task alone also failed to find any periods of significantly greater activity in the two-word composition condition compared with the one-word composition condition in either region at any time. ns, Nonsignificant; \* $p < 0.05$ , \*\* $p < 0.01$ .

within the composition task also identified two clusters of significantly greater activity in the two-word condition compared with the one-word condition at approximately the same times: 163–270 ms ( $p = 0.0168$ ; average  $\pm$  SD: two-word,  $4.03 \pm 1.65$  nAm; one-word,  $2.92 \pm 1.12$  nAm) and 328–410 ms ( $p = 0.0248$ ; average  $\pm$  SD: two-word,  $3.89 \pm 1.94$  nAm; one-word,  $2.70 \pm 1.17$  nAm). Within the list task, no significant clusters of increased two-word activity were identified at any time in the RATL (all clusters  $p > 0.40$ ).

Although these results certainly suggest a role for the RATL within the present manipulation, the exact relation between these effects and combinatorial processing is not quite as clear as in the LATL. Within the LATL, activity for the two-word composition condition increased relative to the other three conditions during the identified cluster. Conversely, within the RATL, the one-word composition condition appears to be the outlier, falling well below the other three conditions for the majority of the interval. Although it is difficult to draw firm conclusions comparing across tasks (because they were performed sequentially and not concurrently), it seems more likely that composition-related activity would display the pattern seen within the LATL (i.e., increased activity in the two-word composition condition relative to the other three conditions). Consequently, the profile within the RATL seems to suggest a suppression of activity within the one-word composition condition rather than an increase in the two-word condition. This interpretation is strengthened by the fact that, in our paradigm, across both tasks the two one-word conditions were visually identical and ostensibly required the same computational demands (i.e., comparing the colored shape with the single object-denoting noun that preceded it). Therefore, the difference in activity levels observed within the RATL suggests a task-related decrease in activity during the one-word composition condition, although it is unclear exactly which aspect of the task might engender this response.

Generally, our results suggest that activity in both the LATL and RATL is modulated by basic combinatorial processes. Specifically, the overall profiles of the two activities suggest that the RATL is subject to a task-related suppression of activity during the one-word composition condition, whereas the LATL reflects basic combinatorial operations  $\sim 225$  ms after the presentation of a composable noun.



**Figure 5.** Full-brain analyses. Plotted regions denote the difference in average amplitude between two-word and one-word conditions for all space time regions in which two-word activity was reliably greater than one-word activity ( $p < 0.05$ , uncorrected) for at least 10 ms over 10 spatial neighbors. For clarity, non-cortical sources have been removed. Results within the composition task (**A**) conform to our ROI analyses and reveal a clear LTL effect from 150 to 250 ms and a clear vmPFC effect from 300 to 450 ms. A sustained difference in activity can also be seen in the RATL from 100 to 400 ms. Additionally, increased activity was also observed within the left superior parietal lobe throughout much of the time interval. Within the list task (**B**), relatively modest increases in two-word activity were seen in the right inferior frontal gyrus and left inferior temporal lobe from  $\sim 200$  to 300 ms.

### Ventromedial prefrontal cortex ROI

Within the vmPFC ROI, the interaction permutation test identified a highly significant, long-lasting cluster of time points for which localized activity exhibited the combinatorial profile (331–480 ms;  $p = 0.0135$ ; average  $\pm$  SD: two-word composition,  $3.61 \pm 2.59$  nAm; one-word composition,  $2.74 \pm 1.57$  nAm; two-word list,  $2.69 \pm 0.95$  nAm; one-word list,  $3.02 \pm 1.71$  nAm) (Fig. 4). A follow-up test within the composition task again conformed to this result, identifying increased activity in the two-word condition compared with the one-word condition from 326 to 442 ms ( $p = 0.0263$ ; average  $\pm$  SD: two-word,  $3.76 \pm 2.95$  nAm; one-word,  $2.78 \pm 1.64$  nAm). Within the list task, no clusters of increased activity for the two-word condition were identified at any point (all clusters  $p > 0.85$ ). Overall, the profile of activity within the vmPFC resembled that observed within the LTL. Activity in the two-word composition condition was clearly greater than in the other three control conditions. This result suggests that activity localized to the vmPFC reflects basic linguistic combinatorial operations  $\sim 300$ –500 ms after the onset of a composable noun.

### Lateral inferior frontal gyrus and posterior temporal lobe ROIs

Within both the LIFG and the LPTL ROIs, the interaction permutation test failed to reveal any significant clusters of combinatorial activity (all clusters  $p = 1$ ) (Fig. 4). A targeted test within the composition task alone also failed to produce any clusters for which activity was significantly greater in the two-word condition compared with the one-word condition in either region (LIFG, all clusters  $p > 0.60$ ; LPTL, all clusters  $p > 0.30$ ). Within the list task, the LIFG showed no significant increases in two-word activity (all clusters  $p > 0.40$ ), whereas the LPTL did exhibit one cluster in which there was a slight trend toward increased activity in the two-word condition (206–266 ms;  $p = 0.1430$ ; average  $\pm$  SD: two-word,  $2.82 \pm 1.66$  nAm; one-word,  $2.17 \pm 1.13$  nAm). This cluster is discussed within the context of the full-brain analysis below.

A division of the LPTL into two ROIs, one encompassing approximately the pMTG and one centered within the AG, also failed to produce any significant clusters of combinatorial activity

(pMTG, all clusters  $p = 1$ ; AG, all clusters  $p > 0.20$ ). However, results from *post hoc* tests within each task do begin to reveal an interesting hint of a dichotomy between these two areas. Activity within the AG demonstrated a trend toward increased two-word activity within the composition task (144–213 ms;  $p = 0.1183$ ; average  $\pm$  SD: two-word,  $2.60 \pm 1.40$  nAm; one-word,  $2.01 \pm 1.01$  nAm) but not within the list task (all clusters  $p > 0.30$ ). Conversely, the pMTG exhibited no combinatorial increases within the composition task (all clusters  $p > 0.55$ ), whereas within the list task, several clusters trended toward increased activity in the two-word condition (107–175 ms,  $p = 0.1141$ ; average  $\pm$  SD: two-word,  $2.72 \pm 1.30$  nAm; one-word,  $2.16 \pm 1.07$  nAm; and 220–273 ms,  $p = 0.1254$ ; two-word,  $3.20 \pm 1.95$  nAm; one-word,  $2.29 \pm 1.12$  nAm). Thus, our results potentially implicate the AG during combinatorial processing, whereas the pMTG seems to play more of a role within the list task. However, because of the extremely tentative nature of these effects, we refrain from additional speculation.

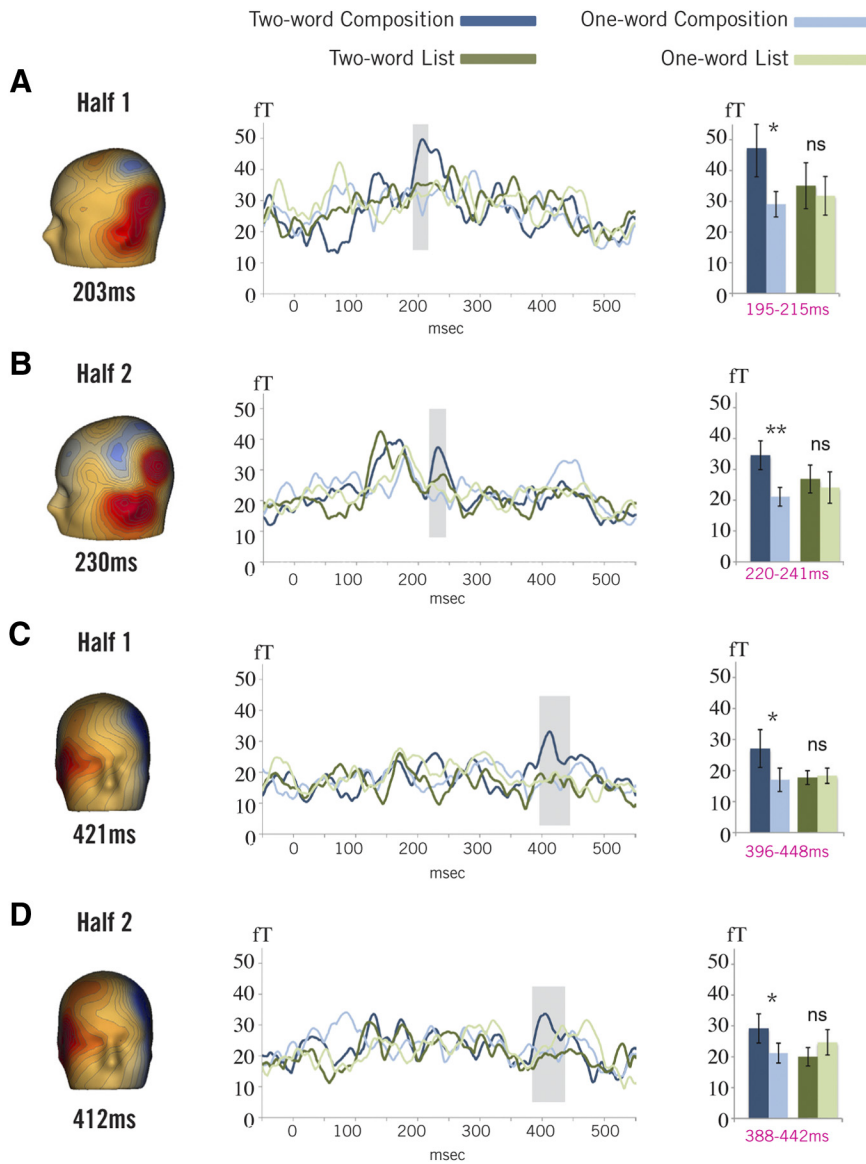
In general, the results from these two areas suggest that past effects observed within the LIFG and LPTL primarily reflect either non-combinatorial or more complex combinatorial mechanisms within language processing. However, it is, of course, also possible that either our paradigm (i.e., MEG in conjunction with our analysis method) is ill-suited to reproduce past results in these regions or the present manipulation is too subtle to adequately bring out combinatorial effects in these regions.

### Full-brain comparisons

In general, our full brain analyses (Fig. 5) conform quite closely to our ROI analyses. Within the composition task, a clear LTL effect can be seen at  $\sim 225$  ms, followed by an increase in two-word activity centered in the vmPFC at  $\sim 400$  ms. Throughout, a sustained difference can also be seen in the RATL, mirroring our previous ROI analysis. No obvious effects appear in either the LIFG or LPTL ROIs.

In the composition task, the only other significant effect of note is a sustained increase in activity during the two-word condition localized to a large region of the left parietal lobe, approximately near the central sulcus and superior parietal lobule (SPL). The central sulcus, of course, houses the primary motor cortices,





**Figure 6.** Sensor analysis results. Each row corresponds to one test from the split-half cross-validation analysis. Waveforms represent the RMS of the sensors corresponding to the dominant field patterns present during the LATL (**A, B**) and vmPFC (**C, D**) source space effects. Shaded regions denote significant clusters of combinatorial activity as identified by a cluster-based permutation test (Maris and Oostenveld, 2007). The field pattern from the grand average of the two-word composition condition is shown for each significant cluster. Within the left hemisphere sensors, the first data half (**A**) produced a significant cluster of combinatorial activity from 195 to 215 ms ( $p = 0.036$ ), within which two-word composition activity was significantly greater than one-word composition activity, and activity in the list conditions did not differ. In the second data half (**B**), a significant cluster of combinatorial activity was identified from 220 to 241 ms ( $p = 0.013$ ) in which two-word composition activity was significantly greater than one-word composition activity, and activity in the list conditions did not differ. Within the frontal sensors, the analysis of the first data half (**C**) revealed a cluster of combinatorial activity from 396 to 448 ms ( $p = 0.011$ ). In this time period, sensor activity was again significantly greater for the two-word composition condition compared with the one-word composition condition. No differences in activity were observed in the list task. In the second data half analysis (**D**), a significant cluster of combinatorial activity was found from 388 to 442 ms ( $p = 0.050$ ). Again, two-word composition activity was significantly greater than one-word composition activity within this cluster, whereas activity in the list conditions showed no significant differences. ns, Nonsignificant; \* $p < 0.05$ ; \*\* $p < 0.01$ .

and activity in this region has been shown to occur not only in overt speech but covert speech as well (Riecker et al., 2000; Huang et al., 2002; Pulvermüller et al., 2006). Thus, it is possible that participants were disproportionately engaged in covert speech during compositional processing compared with non-compositional processing. Another, potentially more intriguing hypothesis arises from past work implicating the SPL in the binding of visual features. Lesions in

the SPL have been associated with deficits in correctly combining the visual features of an object (Friedman-Hill et al., 1995) and also deficits in visual search tasks requiring the conjunction of features (Robertson et al., 1997). Furthermore, increased activity in this region has been observed during conjunction searches compared with those that target only one visual feature (Corbetta et al., 1995). Thus, it is possible that the increase in activity we observe near the SPL is related in some manner to the feature binding required by the two-word composition condition, although the exact nature of this link is unclear at the present time. It should be noted, however, that there is much debate surrounding the role of the SPL in visual feature binding. Many studies suggest that effects in this region during visual conjunction search are primarily attributable not to the binding of features but to the deployment of spatial attention (Donner et al., 2002; Shafritz et al., 2002; Nobre et al., 2003), argued to be a necessary precursor to visual feature binding (Treisman and Gelade, 1980). Within the context of this hypothesis, the present result might be related to a preemptive deployment of attention in preparation for the upcoming target shape in the two-word composition condition because participants are aware binding will be needed for their forthcoming decision given the preceding adjective. However, although parallels between the present task and past visual search studies are intriguing, hypotheses relating the present results to visual attention or feature binding can only be speculative at the present time.

Within the list task, increased activity for the two-word compared with the one-word condition can be seen within the inferior, posterior portion of the left temporal lobe and within the right inferior frontal gyrus (RIFG) at ~200–300 ms. The RIFG has long been implicated in verbal working memory (Petrides et al., 1993; Cohen et al., 1994; Fiez et al., 1996; Wei et al., 2004). Thus, it is not surprising to find increased activity within this region during the present task when participants had to remember two words compared with one. However, not only have past effects generally been found bilaterally, but activity within the LIFG is usually greater than in the RIFG during memory tasks (Paulesu et al., 1993; Awh et al., 1996; Rypma et al., 1999). Thus, the appearance of only a right hemisphere effect within the present contrast is somewhat unexpected. The left temporal effect, which drives the trend toward increased two-word list activity observed in the LPTL ROI, is also somewhat unexpected because this area has not canonically been linked to verbal memory tasks. However, increased activity in this general region has been reported

during nonverbal memory tasks (Courtney et al., 1996; Ungerleider et al., 1998).

Compared with past experiments, the manipulation we used in the list task is relatively weak with respect to memory demands. Thus, the effects that we observed might only make up a muted and incomplete snapshot of the full network of neural regions that participate in this task. Importantly, however, the manipulation within the list task with respect to the number of lexical items in each condition is equivalent to that in the composition task. Therefore, any effects seen within this critical task attributable merely to the difference in lexical material between conditions should have been observable within the list contrast as well. The only hint of such shared activity is within the RIFG because increased activity in this region can be seen from 200 to 300 ms in the composition contrast as well. However, no vmPFC or LATL effects are visible at all within the full-brain list comparison. This suggests that the effects observed within these regions during the composition task are not attributable simply to a difference in the number of lexical items.

### Sensor space analysis

The two clearest combinatorial effects resulting from our source space analysis were in the LATL from ~175 to 275 ms and in the vmPFC from ~300 to 500 ms. Thus, we performed two separate sensor space analyses targeted at further characterizing the relationship between these source space effects and the recorded sensor data.

### Left anterior sensors

Both split-half analyses revealed a robust combinatorial effect within the left hemisphere sensors during the same time window as the LATL source space effect, ~175–275 ms (Fig. 6). Across both halves, the dominant field pattern observed within the grand average of the two-word composition condition at this time and location could be identified in 18 of the 20 participants; however, in each half, a different set of two participants failed to show this component. In the first half, within the targeted time window from 175 to 275 ms, the interaction permutation test revealed a significant cluster of combinatorial activity from 195 to 215 ms ( $p = 0.036$ ; average  $\pm$  SD: two-word composition,  $48.24 \pm 40.60$  fT; one-word composition,  $28.26 \pm 18.58$  fT; two-word list,  $35.15 \pm 31.78$  fT; one-word list,  $31.80 \pm 28.27$  fT). A follow-up test within the composition task identified a significant cluster of increased activity in the two-word condition compared with the one-word condition from 196 to 235 ms ( $p = 0.037$ ; average  $\pm$  SD: two-word,  $46.78 \pm 39.35$  fT; one-word,  $30.65 \pm 20.66$  fT). No such clusters were found within the list task (all clusters  $p > 0.25$ ). In the expanded time window of 0–500 ms, the significance level of these clusters was reduced, reflecting the loss in power attributable to partitioning the data (interaction,  $p = 0.19$ ; composition,  $p = 0.17$ ). In the second data half, the targeted permutation tests again found both a significant cluster of combinatorial activity using the interaction test (220–241 ms;  $p = 0.013$ ; average  $\pm$  SD: two-word composition,  $34.89 \pm 20.10$  fT; one-word composition,  $21.17 \pm 12.91$  fT; two-word list,  $27.12 \pm 19.22$  fT; one-word list,  $24.09 \pm 21.72$  fT) and a significant cluster of increased two-word activity in the composition task (218–255 ms;  $p = 0.018$ ; average  $\pm$  SD: two-word,  $32.41 \pm 17.62$  fT; one-word,  $20.28 \pm 11.19$  fT). No significant effects were identified in the list task (all clusters  $p > 0.8$ ). In the expanded time window, the significance of the identified clusters was again somewhat reduced (interaction,  $p = 0.08$ ; composition,  $p = 0.12$ ).

### Anterior midline sensors

Both split-half analyses also revealed a robust combinatorial effect within the frontal sensors during the later time window surrounding the identified vmPFC source effect, 300–500 ms (Fig. 6). In the first test set, three participants failed to exhibit the dominant frontal field pattern, whereas in the second set, two participants were excluded, with one participant overlapping between the two exclusion sets. In the first split-half analysis, a targeted interaction permutation test from 300 to 500 ms revealed a significant cluster of combinatorial activity from 396 to 448 ms ( $p = 0.011$ ; average  $\pm$  SD: two-word composition,  $27.10 \pm 25.12$  fT; one-word composition,  $17.12 \pm 15.64$  fT; two-word list,  $17.68 \pm 9.18$  fT; one-word list,  $18.35 \pm 10.84$  fT). A follow-up permutation test within the composition task showed a significant cluster of increased two-word activity from 371 to 448 ms ( $p = 0.012$ ; average  $\pm$  SD: two-word,  $25.51 \pm 21.77$ ; one-word,  $16.00 \pm 13.77$  fT). No such clusters were identified within the list task (all clusters  $p > 0.40$ ). The relaxed time window from 0 to 500 ms did not qualitatively reduce the significance of these results (interaction,  $p = 0.027$ ; composition,  $p = 0.030$ ). In the second data half, the targeted permutation tests revealed a significant cluster of combinatorial activity from 388 to 442 ms ( $p = 0.050$ ; average  $\pm$  SD: two-word composition,  $29.20 \pm 20.01$  fT; one-word composition,  $21.24 \pm 13.80$  fT; two-word list,  $20.11 \pm 12.65$  fT; one-word list,  $24.80 \pm 17.43$  fT) and a significant cluster of increased two-word activity within the composition task from 389 to 439 ms ( $p = 0.033$ ; average  $\pm$  SD: two-word,  $29.38 \pm 20.51$  fT; one-word,  $21.02 \pm 13.65$  fT). No significant cluster of increased two-word activity was found within the list task (all clusters  $p > 0.85$ ). This time, expanding the time window slightly reduced the significance values of the two identified clusters (interaction,  $p = 0.11$ ; composition,  $p = 0.15$ ).

Thus, the results of our sensor space analysis indicate that both the LATL and vmPFC source space effects are reflected rather robustly in sensor space as well. A significant combinatorial effect was found in the left hemisphere sensors in the same time range as the LATL effect, ~175–275 ms. Likewise, an analysis of the frontal sensors revealed a significant cluster of combinatorial activity in the same time window as the vmPFC effect, ~300–500 ms. The dominant field pattern associated with this latter effect appears to be the AMF (Fig. 6), thus replicating previous findings relating this field pattern to vmPFC source activity during the resolution of semantic mismatches (Pykkänen and McElree, 2007).

### Discussion

The present study deviates from the neurolinguistic norm of treating basic combinatorial processing as the baseline against which more complex operations are measured. Instead, we explicitly investigated the most minimal of combinatorial processes: the simple composition of an adjective with a noun. Using MEG, we isolated neural activity elicited by the comprehension of a simple, object-denoting noun (“boat”) and measured increases in this activity evoked within a minimal compositional context (“red boat”) compared with a matched control (“xkq boat”). By contrasting this manipulation to that of a similar non-compositional context (“cup, boat”) compared with the same control, we identified activity associated with basic composition that cannot be attributed simply to an increase in lexical–semantic material. Although several past studies have compared combinatorial and non-combinatorial linguistic processing, none have reduced the scope of their investigation to such a minimal and fundamental level. Furthermore, because these past studies

have used either fMRI or PET, the present results provide a unique temporal map of these critical operations.

### ATL and vmPFC effects of composition

Both the LATL and vmPFC demonstrated significant combinatorial activity during our minimal manipulation. In both regions, we found clear clusters of time points for which activity during the processing of the nouns was significantly greater during the compositional context compared with the matched control. The LATL effect peaked at ~225 ms and was followed by increased vmPFC activity at ~400 ms. Importantly, no corresponding increases were observed in the non-compositional contrast. The RATL also exhibited a clear difference in activity during the compositional manipulation, but the activity profile in this region instead suggests a task-related suppression during the control condition because activity in this condition was significantly lower than in the other, structurally identical one-word control.

Both the LATL and RATL have been implicated previously in sentence processing. The comprehension of sentences consistently produces increased activity in these regions compared with unstructured word lists (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000). Thus, the present results suggest that very basic combinatorial operations drive the LATL effects observed in these manipulations. Regarding the RATL, our results are more ambivalent and introduce the possibility that past “increases” observed within this region might in fact reflect decreases in neural activity during non-compositional processing. Although variability in experimental design makes direct comparisons difficult, the markedly different activity profiles exhibited by the LATL and RATL in the present results may help to illuminate past asymmetries found between these two regions. In previous studies, RATL effects are consistently smaller than LATL effects (Mazoyer et al., 1993; Stowe et al., 1998) and occasionally not present at all (Bottini et al., 1994; Vandenberghe et al., 2002). Whether the dichotomy suggested here—that the LATL subserves basic linguistic composition whereas the RATL undergoes a task-related suppression—can explain these differences will require more work.

The vmPFC has been implicated recently as important in resolving semantic mismatches (Pylkkänen and McElree, 2007; Brennan and Pylkkänen, 2008). Several studies have compared compositionally transparent constructions with syntactically matched controls that require more work to obtain a coherent meaning. Results have consistently implicated the vmPFC as important in resolving the more semantically intensive expressions. The present findings, therefore, suggest that the vmPFC supports the construction of basic linguistic meaning and that past effects were not driven by more specialized mechanisms related to semantic mismatch resolution in particular.

This hypothesis—that the vmPFC reflects basic semantic composition—raises an intriguing possibility within the context of the present results. All linguistic combination, even that associated with basic adjective–noun phrases, can be broken down into two broad types of combinatorial processing, the construction of phrases based on grammatical categories (syntactic composition) and the creation of complex meanings from simpler pieces (semantic composition). In contrast to the vmPFC, the LATL has been most strongly associated with syntactic combinatorial processing. Activity in the LATL correlates with measures of syntactic complexity during natural story comprehension (Brennan et al., 2010) and has been found to exhibit reduced activity in contexts that elicit syntactic priming (Noppeney and Price, 2004). Therefore, although our design did not aim to explicitly

disentangle these two types of combinatorial processes, the present results are consistent with the interpretation that the LATL subserves basic syntactic combinatorial operations, whereas the vmPFC supports basic semantic composition. Furthermore, a broad range of processing models mirror the temporal ordering of our results and place the initiation of syntactic composition before semantic composition (Friederici, 2002).

Of course, neither this hypothesis nor our results imply that the LATL and vmPFC solely perform combinatorial operations. Both regions have also been implicated in non-combinatorial, lexical-level tasks as well, such as lexical decision (Nobre et al., 1994; Mummery et al., 1999; Fujimaki et al., 2009), repetition detection (Halgren et al., 1994), and semantic relatedness judgments (Vandenberghe et al., 1996). Additionally, both regions exhibit classical lexical-level effects, such as decreased activity for more frequent (Halgren et al., 2002) or primed (Mummery et al., 1999; Dale et al., 2000; Dhond et al., 2001; Marinkovic et al., 2003) words. However, in the present study, lexical computations should be relatively constant across both tasks. The same words served as critical items in each condition, and the probability of a particular critical item after any preceding item was equated across all stimulus pairs within the context of the experiment. Therefore, although our results clearly do not deny a role for the vmPFC or LATL in non-combinatorial, lexical-level processing, it is not immediately apparent how such operations would produce the effects observed in the present experiment.

It should be emphasized that the partitioning of syntactic and semantic combinatorial processes to the LATL and vmPFC can at best be tentative at the present time. In general, disentangling syntactic and semantic processes is a very difficult problem, with disagreement on even the extent to which a solution is theoretically possible (Pylkkänen, 2008). The majority of studies investigating compositional processes, including the present one, manipulate syntax and semantics in tandem, thus making it difficult to assign either one process or the other to observed effects, and past attempts to disentangle the two have not always produced converging evidence. In particular, although much evidence implicates the LATL in syntactic combinatorial processing, a number of researchers have also suggested that this region plays an important role in semantic combinatorial operations as well. For example, when participants are asked to monitor sentences for semantic anomalies, a subset of the LATL shows increased activity compared with monitoring for syntactic violations (Rogalsky and Hickok, 2009). Also, scrambling the content words of sentences modulates activity in the LATL, although the direction of this effect has varied, with scrambled sentences eliciting both more (Vandenberghe et al., 2002) and less (Humphries et al., 2006) LATL activity compared with normal sentences. Clearly, much work is still needed before any conclusions can be solidly established regarding the apportionment of syntactic and semantic combinatorial operations to specific cortical regions, but the present study does provide a novel, minimal framework on which to build, and the proposed delineation provides an intriguing point of departure for future studies.

### Lack of LIFG and LPTL effects

Neither the LIFG nor LPTL exhibited any significant activity related to basic combinatorial processing. Of course, as with any null result, the absence of these effects may be attributable to any number of factors, such as the particular technique used (the majority of past studies implicating these areas in language processing were conducted within fMRI, whereas the present study localizes MEG data), the analysis method (perhaps a more nu-

anced localization method might provide more power), or the subtlety of the manipulation (the simple act of composing an adjective with a noun might not be enough to drive certain types of combinatorial activity sufficiently above the background noise). However, these null results are also consistent with past evidence suggesting that these regions support either relatively complex or non-combinatorial (i.e., lexical-level) linguistic operations.

Within the LPTL, the pMTG has primarily been linked to long-term lexical storage and access (Pykkänen et al., 2002; Gold et al., 2006). If this interpretation is correct, then one would not expect to see effects in this region during the present manipulation because, as discussed above, lexical computations should be constant across tasks. The more posterior portion of the LPTL, the AG, has shown increased activity for semantically coherent sentences compared with syntactically well-formed but incoherent sentences (Humphries et al., 2007). However, this region, like the pMTG, has also primarily been associated with lexical-level tasks (Rissman et al., 2003; Binder et al., 2005). Thus, the lack of any strong combinatorial effects within the LPTL in the present study is consistent with this region subserving either lexical-level operations or more complex aspects of sentence processing not elicited by the present manipulation.

The LIFG has canonically been associated with the processing of complex syntactic structures (Stromswold et al., 1996) and high-level cognitive functioning (Badre and Wagner, 2007). Thus, we do not find the lack of combinatorial effects within this region particularly surprising given the present minimal manipulation. This outcome would be predicted by any number of hypotheses posited for this region, e.g., syntactic movement (Grodzinsky and Santi, 2008), selection among alternatives (Thompson-Schill et al., 2005), or executive functioning (Koechlin and Summerfield, 2007). Therefore, although no firm conclusions can be drawn from the absence of effects in the LIFG and LPTL, these null results do lend force to the increasing amount of data suggesting that the traditional neurophysiological model of language, revolving around Broca's and Wernicke's areas, is neither necessary nor sufficient as a model for even the most basic linguistic processing.

## Conclusion

Surprisingly, direct investigations into the neural underpinnings of basic combinatorial processing in language have been virtually nonexistent. The present paradigm introduces a powerful method for directly investigating these operations by allowing the linguistic expressions under consideration to be reduced to the absolute minimum: a simple adjective composed with a noun. Interestingly, our manipulation did not reveal any increases in activity within the more traditional Broca's or Wernicke's areas during basic combinatorial processing. Thus, our results add to the growing body of evidence that complete neurophysiological models of language must take into consideration a more extended network of neural regions than this simple binodal configuration. Instead, our results indicate that both the LATL and vmPFC play a prominent role in basic linguistic composition. Activity within these regions was significantly greater during composition compared with matched controls, with increased LATL activity preceding increased vmPFC activity during combinatorial processing. Because past work has implicated the LATL in syntactic composition and the vmPFC in semantic composition, these results are consistent with the hypothesis that the LATL supports basic syntactic structure building, whereas the vmPFC subserves the fundamental creation of linguistic mean-

ing. Future work must now be aimed at building on this foundation to construct a more solid and complete understanding of the neural mechanisms that underlie the comprehension and production of more complex linguistic expressions.

## References

- Allen M, Badecker W, Osterhout L (2003) Morphological analysis in sentence processing: an ERP study. *Lang Cognit Process* 18:405–430.
- Awh E, Jonides J, Smith EE, Schumacher EH, Koeppel RA, Katz S (1996) Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. *Psychol Sci* 7:25–31.
- Badre D, Wagner AD (2007) Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 45:2883–2901.
- Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD (2005) Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47:907–918.
- Bavelier D, Corina D, Jezzard P, Padmanabhan S, Clark V, Karni A, Prinster A, Braun A, Lalwani A, Rauschecker J, Turner R, Neville H (1997) Sentence reading: a functional MRI study at 4 tesla. *J Cogn Neurosci* 9:664–686.
- Bever T (1970) The cognitive basis of linguistic structures. In: *Cognition and the development of language* (Hayes J, ed). New York: Wiley.
- Binder JR, Medler DA, Desai R, Conant LL, Liebenthal E (2005) Some neurophysiological constraints on models of word naming. *Neuroimage* 27:677–693.
- Bookheimer S, Zeffiro T, Blaxton T, Gaillard W, Theodore W (1995) Regional cerebral blood flow during object naming and word reading. *Hum Brain Mapp* 3:93–106.
- Bottini G, Corcoran R, Sterzi R, Paulesu E, Schenone P, Scarpa P, Frackowiak RS, Frith CD (1994) The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain* 117:1241–1253.
- Brennan J, Pykkänen L (2008) Processing events: behavioral and neuro-magnetic correlates of aspectual coercion. *Brain Lang* 106:132–143.
- Brennan J, Pykkänen L (2010) Processing psych verbs: behavioral and MEG measures of two different types of semantic complexity. *Lang Cognit Process* 25:777–807.
- Brennan J, Nir Y, Hasson U, Malach R, Heeger DJ, Pykkänen L (2010) Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain Lang*. Advance online publication. Retrieved December 30, 2010. doi:10.1016/j.bandl.2010.04.002.
- Broca P (1861) Perte de la parole. Romolissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bull Soc Anatom* 2:235–238.
- Caplan D, Alpert N, Waters G, Olivieri A (2000) Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Hum Brain Mapp* 9:65–71.
- Cohen J, MacWhinney B, Flatt M, Provost J (1993) PsyScope: an interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behav Res Methods Instr Comp* 25:257–271.
- Cohen JD, Forman SD, Braver TS, Casey B, Servan-Schreiber D, Noll D (1994) Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Hum Brain Mapp* 1:293–304.
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–608.
- Corbetta M, Shulman GL, Miezin FM, Petersen SE (1995) Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science* 270:802–805.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1996) Object and spatial working memory activate separate neural systems in human cortex. *Cereb Cortex* 6:39–49.
- Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26:55–67.
- Dhond RP, Buckner RL, Dale AM, Marinkovic K, Halgren E (2001) Spatio-temporal maps of brain activity underlying word generation and their modification during repetition priming. *J Neurosci* 21:3564–3571.
- Donner TH, Kettermann A, Diesch E, Ostendorf F, Villringer A, Brandt SA (2002) Visual feature and conjunction searches of equal difficulty en-

- gage only partially overlapping frontoparietal networks. *Neuroimage* 15:16–25.
- Embick D, Hackl M, Schaeffer J, Kelepir M, Marantz A (2001) A magnetoencephalographic component whose latency reflects lexical frequency. *Cogn Brain Res* 10:345–348.
- Fiebach CJ, Friederici AD, Smith EE, Swinney D (2007) Lateral inferotemporal cortex maintains conceptual-semantic representations in verbal working memory. *J Cogn Neurosci* 19:2035–2049.
- Fiez JA, Raife EA, Balota DA, Schwarz JP, Raichle ME, Petersen SE (1996) A positron emission tomography study of the short-term maintenance of verbal information. *J Neurosci* 16:808–822.
- Friederici AD (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78–84.
- Friederici AD, Meyer M, von Cramon DY (2000) Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang* 74:289–300.
- Friederici AD, Fiebach CJ, Schlesewsky M, Bornkessel ID, von Cramon DY (2006) Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb Cortex* 16:1709–1717.
- Friedman-Hill SR, Robertson LC, Treisman A (1995) Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science* 269:853–855.
- Fujimaki N, Hayakawa T, Ihara A, Wei Q, Munetsuna S, Terazono Y, Matani A, Murata T (2009) Early neural activation for lexico-semantic access in the left anterior temporal area analyzed by an fMRI-assisted MEG multipole method. *Neuroimage* 44:1093–1102.
- Gitelman DR, Nobre AC, Sonty S, Parrish TB, Mesulam MM (2005) Language network specializations: an analysis with parallel task designs and functional magnetic resonance imaging. *Neuroimage* 26:975–985.
- Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH (2006) Dissociation of automatic and strategic lexical-semantic: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *J Neurosci* 26:6523–6532.
- Grodzinsky Y, Santi A (2008) The battle for Broca's region. *Trends Cogn Sci* 12:474–480.
- Hagoort P, Hald L, Bastiaansen M, Petersson KM (2004) Integration of word meaning and world knowledge in language comprehension. *Science* 304:438–441.
- Halgren E, Baudena P, Heit G, Clarke JM, Marinkovic K, Chauvel P, Clarke M (1994) Spatio-temporal stages in face and word processing. 2. Depth-recorded potentials in the human frontal and Rolandic cortices. *J Physiol (Paris)* 88:51–80.
- Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, Dale AM (2002) N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage* 17:1101–1116.
- Hillebrand A, Barnes GR (2002) A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage* 16:638–650.
- Holcomb PJ (1993) Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology* 30:47–61.
- Huang J, Carr TH, Cao Y (2002) Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum Brain Mapp* 15:39–53.
- Humphries C, Willard K, Buchsbaum B, Hickok G (2001) Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *NeuroReport* 12:1749–1752.
- Humphries C, Love T, Swinney D, Hickok G (2005) Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum Brain Mapp* 26:128–138.
- Humphries C, Binder JR, Medler DA, Liebenthal E (2006) Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J Cogn Neurosci* 18:665–679.
- Humphries C, Binder JR, Medler DA, Liebenthal E (2007) Time course of semantic processes during sentence comprehension: an fMRI study. *Neuroimage* 36:924–932.
- Ischebeck A, Indefrey P, Usui N, Nose I, Hellwig F, Taira M (2004) Reading in a regular orthography: an fMRI study investigating the role of visual familiarity. *J Cogn Neurosci* 16:727–741.
- Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive function. *Trends Cogn Sci* 11:229–235.
- Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207:203–205.
- Lau EF, Phillips C, Poeppel D (2008) A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci* 9:920–933.
- Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E (2003) Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38:487–497.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Mazoyer B, Tzourio N, Frak V, Syrota A, Murayama N, O L, Salamon G, Dehaene S, Cohen L, Mehler J (1993) The cortical representation of speech. *J Cogn Neurosci* 5:467–479.
- Miller G, Chomsky N (1963) Finitary models of language users. In: *Handbook of mathematical psychology* (Luce R, Bush R, Galanter E, eds), pp 419–491. New York: Wiley.
- Mosher JC, Leahy RM (1998) Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE Trans Biomed Eng* 45:1342–1354.
- Mummery CJ, Shallice T, Price CJ (1999) Dual-process model in semantic priming: a functional imaging perspective. *Neuroimage* 9:516–525.
- Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372:260–263.
- Nobre AC, Coull JT, Walsh V, Frith CD (2003) Brain activations during visual search: contributions of search efficiency versus feature binding. *Neuroimage* 18:91–103.
- Noppeney U, Price CJ (2004) An fMRI study of syntactic adaptation. *J Cogn Neurosci* 16:702–713.
- Osterhout L, Holcomb PJ, Swinney DA (1994) Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *J Exp Psychol Learn Mem Cogn* 20:786–803.
- Paulesu E, Frith CD, Frackowiak RS (1993) The neural correlates of the verbal component of working memory. *Nature* 362:342–345.
- Petrides M, Alivisatos B, Meyer E, Evans AC (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci U S A* 90:878–882.
- Potter MC, Faulconer BA (1979) Understanding noun phrases. *J Verbal Learn Verbal Behav* 18:509–521.
- Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, Shtyrov Y (2006) Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci U S A* 103:7865–7870.
- Pykkänen L (2008) Mismatching meanings in brain and behavior. *Lang Ling Compass* 2:712–738.
- Pykkänen L, Marantz A (2003) Tracking the time course of word recognition with MEG. *Trends Cogn Sci* 7:187–189.
- Pykkänen L, McElree B (2007) An MEG study of silent meaning. *J Cogn Neurosci* 19:1905–1921.
- Pykkänen L, Stringfellow A, Flagg E, Marantz A (2000) A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. In: *Proceedings of the 12th International Conference on Biomagnetism*, Helsinki University of Technology, Espoo, Finland, August.
- Pykkänen L, Stringfellow A, Marantz A (2002) Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain Lang* 81:666–678.
- Pykkänen L, Oliveri B, Smart AJ (2009a) Semantics vs. world knowledge in prefrontal cortex. *Lang Cognit Process* 24:1313–1334.
- Pykkänen L, Martin AE, McElree B, Smart A (2009b) The anterior midline field: coercion or decision making? *Brain Lang* 108:184–190.
- Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W (2000) Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport* 11:1997–2000.
- Rissman J, Eliassen JC, Blumstein SE (2003) An event-related fMRI investigation of implicit semantic priming. *J Cogn Neurosci* 15:1160–1175.
- Robertson L, Treisman A, Friedman-Hill S, Grabowecy M (1997) The interaction of spatial and object pathways: evidence from Balint's Syndrome. *J Cogn Neurosci* 9:295–317.
- Rogalsky C, Hickok G (2009) Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cereb Cortex* 19:786–796.
- Rugg MD (1985) The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology* 22:642–647.
- Rypma B, Prabhakaran V, Desmond JE, Glover GH, Gabrieli JD (1999)

- Load-dependent roles of frontal brain regions in the maintenance of working memory. *Neuroimage* 9:216–226.
- Salmelin R (2007) Clinical neurophysiology of language: the MEG approach. *Clin Neurophysiol* 118:237–254.
- Santi A, Grodzinsky Y (2007) Taxing working memory with syntax: bihemispheric modulations. *Hum Brain Mapp* 28:1089–1097.
- Shafritz KM, Gore JC, Marois R (2002) The role of the parietal cortex in visual feature binding. *Proc Natl Acad Sci U S A* 99:10917–10922.
- Sternberg S (1967) Retrieval of contextual information from memory. *Psychonom Sci* 8:55–56.
- Stowe LA, Broere CA, Paans AM, Wijers AA, Mulder G, Vaalburg W, Zwarts F (1998) Localizing components of a complex task: sentence processing and working memory. *NeuroReport* 9:2995–2999.
- Stromswold K, Caplan D, Alpert N, Rauch S (1996) Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 52:452–473.
- Tarkiainen A, Helenius P, Hansen PC, Cornelissen PL, Salmelin R (1999) Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122:2119–2132.
- Thompson-Schill SL, Bedny M, Goldberg RF (2005) The frontal lobes and the regulation of mental activity. *Curr Opin Neurobiol* 15:219–224.
- Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136.
- Ungerleider LG, Courtney SM, Haxby JV (1998) A neural system for human visual working memory. *Proc Natl Acad Sci U S A* 95:883–890.
- Van Petten C, Kutas M (1990) Interactions between sentence context and word frequency in event-related brain potentials. *Mem Cognit* 18:380–393.
- Vandenberghe R, Price C, Wise R, Josephs O, Frackowiak RS (1996) Functional anatomy of a common semantic system for words and pictures. *Nature* 383:254–256.
- Vandenberghe R, Nobre AC, Price CJ (2002) The response of left temporal cortex to sentences. *J Cogn Neurosci* 14:550–560.
- Wei X, Yoo SS, Dickey CC, Zou KH, Guttmann CR, Panych LP (2004) Functional MRI of auditory verbal working memory: long-term reproducibility analysis. *Neuroimage* 21:1000–1008.
- Wernicke C (1874) *Der aphasische symptom-complex: eine psychologische studie auf anatomischer basis*. Breslau, Poland: Cohn and Weigert.