Basic Linguistic Composition Recruits the Left Anterior Temporal Lobe and Left Angular Gyrus During Both Listening and Reading

D.K. Bemis[1] and L. Pylkkänen[1,2]

[1]Department of Psychology and [2]Department of Linguistics, New York University, New York, NY, USA

Address correspondence to Douglas K. Bemis, New York University, 6 Washington Pl., New York, NY 10003, USA. Email: doug.bemis@nyu.edu

Language is experienced primarily through one of two mediums— spoken words and written text. Although substantially different in form, these two linguistic vehicles possess similar powers of expression. Consequently, one goal for the cognitive neuroscience of language is to determine where, if anywhere, along the neural path from sensory stimulation to ultimate comprehension these two processing streams converge. In the present study, we investigate the relationship between basic combinatorial operations in both reading and listening. Using magnetoencephalography, we measured neural activity elicited by the comprehension of simple adjective– noun phrases (*red boat*) using the same linguistic materials and tasks in both modalities. The present paradigm deviates from previous cross-modality studies by investigating only basic combinatorial mechanisms—specifically, those evoked by the construction of simple adjective–noun phrases. Our results indicate that both modalities rely upon shared neural mechanisms localized to the left anterior temporal lobe (lATL) and left angular gyrus (lAG) during such processing. Furthermore, we found that combinatorial mechanisms subserved by these regions are deployed in the same temporal order within each modality, with lATL activity preceding lAG activity. Modality-specific combinatorial effects were identified during initial perceptual processing, suggesting top-down modulation of low-level mechanisms even during basic composition.

Keywords: auditory/visual, basic combinatorics, language, magnetoencephalography, minimal phrases

Physically, linguistic expressions primarily impinge upon the human consciousness in one of two ways, through speech— consisting of vibrations in the air—or reading—conveyed through photons reflecting off of a physical surface. Despite these initial differences, both methods of expression ultimately give rise to similar mental states in the mind of a comprehender. While there have been many investigations into the convergence of these input streams during single word comprehension (e.g. Marinkovic et al. 2003) and several recent studies comparing the processing of complex linguistic expressions across modalities (e.g. Jobard et al. 2007), the extent to which basic combinatorial mechanisms in language are shared across modalities is unknown. Although past results generally support the view that speech and reading rely upon a common set of neural mechanisms, systematic discrepancies between neural effects generated by auditory and visual processing of complex expressions (e.g. Michael et al. 2001) make it unclear whether past similarities reflect shared basic combinatorial mechanisms or a common reliance on more complex processes such as working memory and cognitive control. Disentangling these possibilities is vital for determining the extent to which results acquired in one domain of language processing can be applied to the other; a question especially important within the fields of psycho- and neuro-linguistics as the majority of linguistic experiments have been carried out within the visual modality alone.

In the present work, we assess the extent to which basic linguistic combinatorial operations are shared between auditory and visual processing. Using a minimal magnetoencephalography (MEG) paradigm, we measured neural activity associated with the comprehension of a simple noun (e.g. *boat*) and identified increases in this activity when the noun was combined with a preceding adjective (e.g. *red*) compared with when it was preceded by a non-word (e.g. *xbl*) or a non-compositional list control (e.g. *cup*). Within each subject, combinatorial processing was measured in both the auditory and visual modalities using the same task and linguistic materials. Thus, we were able to search for evidence of shared combinatorial mechanisms by comparing between the two sets of results and identifying cortical sources that were preferentially active during basic linguistic combination regardless of modality.

Sensory Level Operations

Clearly, given the vast differences in the physical nature of speech and written words, initial sensory processing of the two must diverge to a large degree. During reading, a visual word must first be broken down into basic visual units such as lines and intersections. These units must then be constructed into increasingly complex visual representations such as letters, graphemes and finally an orthographic representation of the word (see Dehaene 2009). Similarly, when listening to a spoken word, linguistic input must first be analyzed into basic auditory units such as different frequency bands. These features must then be combined into progressively larger and more complex segments such as phonemes, syllables, and, eventually, phonetic representations of words (see Hickok and Poeppel 2007). Unsurprisingly, these different functional demands give rise to consistently different neuroimaging results between the two modalities. In general, although many cortical regions are active during both visual and auditory language processing, reading single words shows increased activity in bilateral occipital and inferior temporal regions compared with listening, which in turn evokes more activity bilaterally in the superior temporal gyri (e.g. Petersen et al. 1988; Booth et al. 2002, 2003). Temporally, there is strong evidence from electrophysiological studies that these differential regions are involved in early processing stages, generally completing around 200 ms following the presentation of the stimulus.

MEG measures of estimated neural activity indicate that, prior to 200 ms, visually presented words primarily evoke activity in the medial, posterior occipital cortex followed by the left inferior temporal cortex (see Salmelin 2007). Initial occipital activity peaks at approximately 100 ms following

word onset and has been shown to reflect the basic analysis of visual features as the amplitude of this neural measure increases with the amount of noise in a visual display and the complexity of a visual stimulus (Tarkiainen et al. 1999; Tarkiainen et al. 2002). The second component localizes to the fusiform gyrus at approximately 150 ms and responds preferentially to letter strings compared with symbols (Tarkiainen et al. 1999; Tarkiainen et al. 2002). This component is hypothesized to reflect the construction of complex visual representations from initial visual features as this activity has also been shown to increase during the processing of complex non-linguistic stimuli such as faces (e.g. Xu et al. 2005) and can be modulated by the orthographic and even morphological properties of words (Zweig and Pylkkänen 2009; Solomyak and Marantz 2010).

During auditory speech comprehension, similar initial computations (i.e. the conversion of raw sensory data into primitive features followed by the construction of complex perceptual representations) are also hypothesized to monopolize processing before 200 ms, however, this activity is localized within the superior temporal cortex as opposed to the posterior occipital and inferior temporal regions (see Salmelin 2007). MEG studies have revealed an initial evoked component that responds to basic acoustic and phonetic properties of auditory words peaking at approximately 100 ms (Kuriki and Murase 1989; Parviainen et al. 2005). This component has been localized to the superior temporal cortex bilaterally and is hypothesized to reflect the initial breakdown of complex auditory stimuli into basic perceptual features (Salmelin 2007). The timing and nature of secondary processing of acoustically complex stimuli has been inferred more indirectly than for visual stimuli, as auditory studies have generally employed mismatch paradigms to investigate early acoustic processing. Results from these studies have found an early component, occurring at approximately 180 ms, that responds to acoustically unexpected events. This component also localizes to the superior temporal cortex bilaterally and has been shown to be sensitive to phonological categories (Phillips et al. 2000) and native compared with non-native vowels (Näätänen et al. 1997). The neural mechanisms supporting the creation of syllable and ultimately whole word phonological forms from these basic features have been difficult to pin down, however, hemodynamic investigations indicate that the superior temporal sulcus is integrally involved in such operations (Hickok and Poeppel 2004; Liebenthal et al. 2005).

Thus, it is clear from both a theoretical and empirical standpoint that the neural processing streams evoked by visual and auditory linguistic items are highly divergent in the initial stages, both in terms of function and location. Consequently, we expect to see the clearest evidence for shared combinatorial mechanisms following this initial processing phase.

Lexical Level Operations

Before combinatorial operations can proceed, however, perceptual representations are first connected to stored lexical meanings that then enter into various syntactic and semantic combinatory operations, ultimately leading to the comprehension of a full linguistic expression. Hemodynamic comparisons between auditory and visual lexical tasks and manipulations, such as semantic priming, repetition, relatedness judgments, and lexical decision, have in general found many shared regions of activity across modalities primarily in the left inferior frontal gyrus (LIFG) (Buckner et al. 2000; Booth et al.

2002; Price et al. 2003), left middle superior temporal cortex (Howard et al. 1992; Booth et al. 2002; Price et al. 2003; Sass et al. 2009), and left inferior temporal cortex (Buckner et al. 2000; Price et al. 2003; Cohen et al. 2004). In terms of timing, lexical access effects within these regions typically appear following early initial sensory processing (e.g. Marinkovic et al. 2003), and a wide variety of lexical tasks and manipulations have produced similar electrophysiological effects within both modalities. For example, word repetition (350 ms—Rugg and Nieto-Vegas 1999), the processing of inflected words (400 ms—Leinonen et al. 2009), semantic priming (400 ms—Anderson and Holcomb 1995; Gomes et al. 1997), linguistic memory probes (560 ms—Kayser et al. 2003), and the processing of semantically related and unrelated word lists (Vartiainen et al. 2009) have all evoked responses with similar topographies and timings across modalities.

These results have led to the predominant view that the initially divergent perceptual pathways converge upon a single mental lexicon (e.g. Caramazza et al. 1990; Price 2000; Marinkovic et al. 2003)—a position sometimes taken as assumed (e.g. Hagoort and Brown 2000b; Spitsyna et al. 2006). However, neurological data also supports an alternate hypothesis in which there are two distinct semantic systems, one for each modality. Several studies have identified patients who lack the ability to access various types of semantic knowledge when processing linguistic input in one modality but not the other (e.g. Warrington 1975; Beauvois 1982; Shallice 1987; Sheridan and Humphreys 1993; Marangolo et al. 2004). Because these deficits are not accompanied by perceptual impairments, and they do not appear to depend upon the type of semantic information required, several researchers have posited the existence of two separate, modality-specific semantic systems to explain the pattern of results (e.g. Shallice 1988; Druks and Shallice 2000), although the necessity of this dichotomy has been questioned (e.g. Caramazza et al. 1990).

Combinatorial Level Operations

While the debate regarding multiple versus unitary semantic systems remains unresolved, it should be made clear that the outcome does not force an answer one way or another to the present question of whether or not basic linguistic combinatorial operations are shared across modalities. Just as divergent perceptual processes might converge upon a shared lexicon, so too might divergent lexicons submit to operations by a single combinatorial engine. Conversely, it is logically possible, though somewhat unparsimonious, that different input modalities may utilize a shared lexicon while still relying on divergent mechanisms to combine these representations. To reiterate Marr's (1982) famous distinction, even if the computational profiles of auditory and visual combinatorial mechanisms are *identical*—the same inputs go in and the same outputs come out—their implementations may be entirely distinct. For example, written words can connect to meaning either by way of a phonological code or orthography alone (Saffran and Marin 1977; Dérouesné and Beauvois 1979). Such dual routes between auditorily and visually based processing might reflect the different developmental timelines of reading and listening—between modalities, common electrophysiological measures evoked during combinatorial linguistic processing exhibit different properties during development (Holcomb et al. 1992)—or they might reflect the

vastly different temporal structure of speech and writing. During speech comprehension, language enters the combinatorial system in the form of fluid auditory stimulation that unfolds continuously in time. The comprehender has little control over the speed or fluency of the input, and thus the combinatorial engine must be able to flexibly adapt on the fly to varying rates of informational input. In fact, even within the auditory modality alone, it has been suggested that multiple neural mechanisms are needed to parse the highly variable temporal properties of continuous speech (e.g. Poeppel 2003). In contrast, a comprehender encounters visual language in a static manner that allows for a higher degree of control over the rate that information is processed by the system.

These temporal differences have led past researchers (e.g. Hagoort and Brown 2000b) to suggest that different combinatorial processes operate within each modality. This hypothesis is bolstered by the finding that altering the structure of visual stimuli to temporally reflect speech (i.e. dynamically unfolding over time) creates significantly more errors in comprehension than when it is presented in a more natural (i.e. static) format (Lee and Newman 2010). Also, semantic priming effects have been shown to increase with stimulus onset asynchrony (SOA) during auditory processing and decrease with longer SOAs in visual processing (Anderson and Holcomb 1995), and temporal anaphora are more difficult to comprehend during speech compared with reading (Jakimik and Glenberg 1990). One possibility is that these effects reflect a deep neural divide in how the brain converts visual and auditory inputs from form to meaning. Perhaps the static nature of writing lends itself to neural mechanisms that can combine representations in parallel, while auditory processing must instead rely upon more serially arranged processes that reflect the temporally extended nature of speech. On the other hand, it is possible that these effects stem instead from differences in complex, domain-general abilities such as working memory and cognitive control that operate in concert with a shared set of basic combinatorial mechanisms to process variable input. At the present time, it is difficult to adjudicate between these two possibilities as no previous studies have directly investigated the extent to which basic combinatorial linguistic mechanisms are shared across modalities.

In general, the vast majority of psycho- and neuro-linguistic investigations contrast the processing of either complex or unexpected linguistic expressions (and often both) with simpler, but still structurally complex, controls. Thus, it is often difficult to confidently disentangle effects associated with basic combinatorial mechanisms from those resulting from more complex processing, such as the maintenance of working memory or repair strategies. Direct neurolinguistic comparisons between auditory and visual combinatorial linguistic processing have overwhelmingly investigated either violation or complexity manipulations and in both cases canonical results from single modality investigations have been reliably replicated across modalities. For example, increased activity in the LIFG during the processing of object-relative compared with simple expressions was observed in functional magnetic resonance imaging (fMRI) for both modalities (Michael et al. 2001; Constable et al. 2004). Semantic priming effects in the LIFG and temporal cortex, previously reported in single-word studies (e.g. Booth et al. 2002), were also observed following the repetition of sentences that overlapped in semantic content (Devauchelle et al. 2009). Within electrophysiology,

both the N400—canonically elicited by sentence final incongruous words (e.g. Kutas and Hillyard 1980)—and the P600 —consistently evoked following syntactically difficult or anomalous constructions (e.g. Osterhout et al. 1994)—have been observed in both modalities within the same set of subjects (N400—Holcomb et al. 1992; Hagoort and Brown 2000b; Balconi and Pozzoli 2005) (P600—Hagoort and Brown 2000a; Balconi and Pozzoli 2005).

While the convergent nature of these results certainly suggests that mechanisms recruited during the processing of complex linguistic expressions are shared across modalities, it is unclear exactly which aspect of such processing these results reflect. Specifically, it is unclear whether or not these results bear directly upon basic combinatorial processing. Both violation and complexity manipulations might produce results related to basic processing, however, there are a myriad of other processes that might differ between conditions in these manipulations as well, such as the amount of working memory, conflict resolution, or executive control required to process the input. These complexities make it difficult to determine if observed commonalities arise from shared basic combinatorial mechanisms or a mutual dependence on more complex abilities. While most studies have highlighted shared effects across modalities, many investigations have also found differences during auditory and visual combinatorial processing both in the topography of electrophysiological effects (e.g. Niznikiewicz et al. 1997; Hagoort and Brown 2000a; Hagoort and Brown 2000b) and the pattern of hemodynamically activated voxels (Michael et al. 2001; Constable et al. 2004; Lindenberg and Scheef 2007; Buchweitz et al. 2009). Auditory effects generally evoke stronger (Holcomb et al. 1992; Hagoort and Brown 2000b; Michael et al. 2001; Spitsyna et al. 2006), more bilateral effects (Hagoort and Brown 2000b; Carpentier et al. 2001; Michael et al. 2001; Constable et al. 2004; Buchweitz et al. 2009) that occasionally onset earlier and extend longer than their visual counterparts (Holcomb and Neville 1990, 1991; Osterhout and Holcomb 1993), leading several researchers to suggest that additional combinatorial mechanisms are invoked during auditory language comprehension compared with reading (e.g. Hagoort and Brown 2000b; Michael et al. 2001; Constable et al. 2004). On the other hand, several studies have also found that the processing of complex linguistic stimuli activates a wider region of the LIFG when presented visually (encompassing both Brodmann's area (BA) 44 and BA 45) than auditorily (BA 45 only) (Carpentier et al. 2001; Constable et al. 2004). Thus, while past comparisons of auditory and visual combinatorial linguistic processing have found relatively consistent patterns of both shared and distinct neural activity, the complex nature of the manipulations used in these comparisons makes it difficult to assign specific functional roles to either.

In general, there has been a noticeable dearth of studies investigating basic combinatorial mechanisms (i.e. those employed during the comprehension of simple linguistic expressions), even within a single modality. In total, we could find only two cross-modal studies that treat simpler linguistic processing as the condition of interest as opposed to the baseline. Both studies contrasted the comprehension of sentences with unstructured lists of words or non-word stimuli, and both sets of results are consistent with the majority of unimodal manipulations of the same type (e.g. Friederici et al.

2000), finding increased activity in the bilateral anterior temporal lobes during sentence comprehension (Jobard et al. 2007; Lindenberg and Scheef 2007). Compared with manipulations of complexity or expectedness, these results more directly suggest that basic combinatorial linguistic mechanisms are shared across modalities, however, this evidence is still somewhat unclear as the processing investigated in these studies (i.e. the comprehension of complete sentences for both and stories in one case) remains relatively complex and potentially involves a large number of diverse operations. Thus, at the present time, while there is a fair amount of evidence suggesting that certain facets of linguistic combinatorial processing may draw upon a common set of neural circuits in both visual and auditory processing, it is currently unknown whether basic combinatorial mechanisms (i.e. those responsible for the construction of minimal linguistic phrases) are shared across modalities.

The Paradigm

The present paradigm was developed precisely in order to target such basic mechanisms (see Bemis and Pylkkänen 2011 and Fig. 1). In our design, the neural activity of interest is that evoked by a simple, object denoting noun (e.g. *boat*), and the fundamental comparison is between the presentation of this noun within a minimal compositional environment (i.e. preceded by a simple color adjective) compared with a non-compositional control (i.e. preceded by either a non-pronounceable consonant string in the visual modality or a burst of noise in the auditory modality). In order to ensure that subjects remained attentive to the stimuli, we asked them to judge whether or not a colored picture following the noun matched the linguistic stimuli that preceded it. In the adjective–noun condition, both the color and shape were required to match while in the 1-word condition only the shape was relevant to the decision. To control for the difference in lexical-semantic material between these two conditions, we also had subjects complete a control task in which the adjectives were replaced by length-matched nouns. In this task, subjects were instructed to determine whether or not the shape of the following picture was denoted by any of the words that preceded it. In all trials, each stimulus was presented sequentially, allowing us to record neural activity evoked at the point of basic combination (i.e. at the nouns following an adjective) compared with physically identical stimuli for which no combination occurred (i.e. nouns following a noun or non-word).

To compare combinatorial processing between modalities, we administered this paradigm twice to each subject, once with auditory linguistic stimuli and once with visual linguistic stimuli, while the tasks and target pictures were held constant. If basic combinatorial mechanisms are shared across modalities, we expect to find cortical regions for which localized MEG activity is greater during the combinatorial condition compared with the control conditions, regardless of the modality of the linguistic stimuli. If, on the other hand, previously identified shared effects reflect more complex processes, then we might expect to see previously observed modality-specific effects augmented in the present comparison.

Materials and Methods

Participants
Twenty-two non-colorblind, native English speakers participated in the study (13 females, average age: 25.6). All had normal or corrected-to-normal vision and gave informed, written consent.
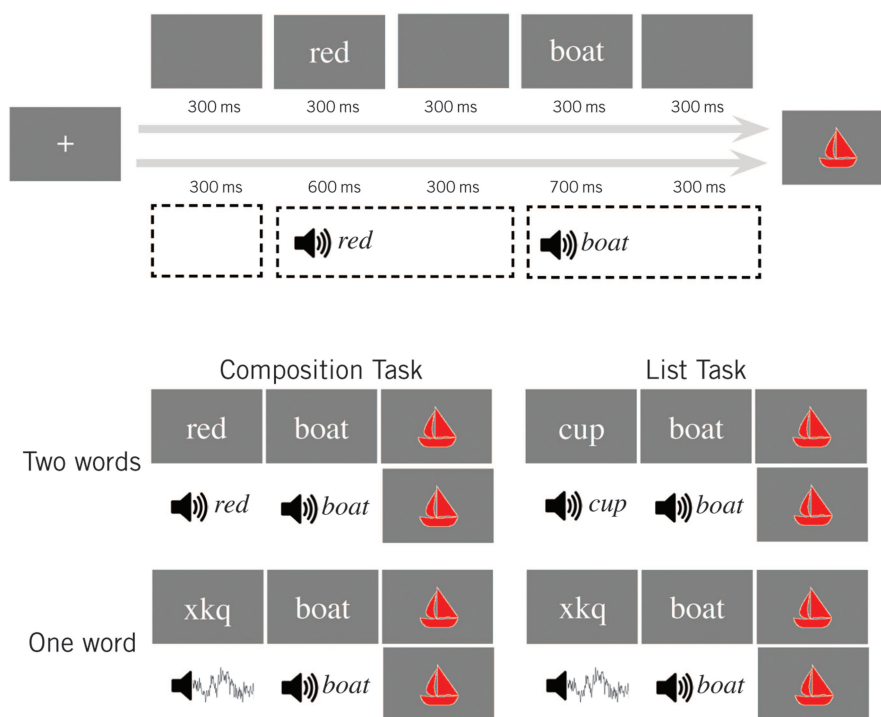

Figure 1. Experimental design. Our design crossed task (composition and list) and number of words (2 and 1). In each trial, participants indicated whether the target picture matched the preceding words. In the composition task, all preceding words were required to match the shape, whereas in the list task, any matching word sufficed. A total of 6 colors and 20 shapes were randomly combined and used as stimuli, and trials were split evenly between matching and non-matching pictures. Only activity recorded at the final word (*boat*) was analyzed. All subjects performed both the auditory and visual versions of the tasks, and the same linguistic materials were used for both sets of tasks.

Visual Stimuli

During the visual tasks (see Fig. 1), each trial contained a small fixation cross, an initial word or non-word, a critical noun, and a target shape. The initial word or non-word varied by condition and was an adjective (2-word Composition condition), noun (2-word List condition), or unpronounceable consonant string (1-word conditions). Twenty one-syllable nouns were employed in the second (critical) stimulus position (*disc, plane, bag, lock, cane, hand, key, shoe, bone, bell, boat, cross, cup, flag, fork, heart, lamp, leaf, star, house*). Adjectives in the 2-word Composition condition (*red, blue, pink, black, green, brown*) were matched in length with the nouns used as first stimuli in the List task (*cane, lamp, disc, fork, bone, heart*). In both tasks, 1-word trials were produced by substituting each adjective or initial noun with a corresponding unpronounceable consonant string of the same length (*xkq, qxsw, mtpv, rjdnw, wvcnz, zbxlv*). Target shapes were manually created to depict canonical, unambiguous representations of the objects denoted by the nouns and were filled in with 1 of the 6 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 Pyschtoolbox (http://psychtoolbox.org/; Brainard 1997; Pelli 1997) and were projected approximately 50 cm from the subject's eye. Linguistic stimuli were displayed in non-proportional Courier font and subtended between 2° and 4° while target shapes were larger, subtending between 6° and 10°.

During each task, subjects viewed 160 trials, 80 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 while the 2-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 20 critical nouns was used 4 times, twice in matching and twice in non-matching trials. Trial and stimuli lists were randomized and constructed separately for each subject.

Auditory Stimuli

The auditory and visual paradigms had identical trial and task structures, and the same stimuli lists were used for both modalities within a single session. The only difference in the auditory paradigm was that the first two stimuli were converted into spoken words (see Fig. 1). Each of the 20 nouns and 6 adjectives used in the visual tasks was recorded by a female speaker using neutral intonation. The average length of the 20 critical nouns was 569 ms (56 ms std.). Two-word auditory conditions used the same 6 adjectives and initial nouns as in the visual paradigm, and these two sets of stimuli did not differ in length ($P > 0.60$, 2-tailed paired *t*-test; adjectives, 540 ms average [28 ms std.]; initial list nouns, 538 ms average [36 ms std.]). In the 1-word conditions, non-pronounceable consonant strings were replaced by pink noise. Pink noise is characterized by a falloff in spectral power inversely proportional to the frequency. This is in contrast to white noise, for which the power is constant across all frequencies and red (or Brownian) noise for which the power decrease is inversely proportional to the square of the frequency. An initial 5-s recording of pink noise (from http://www.mediacollege.com/audio/noise/download.html) was split into 6 distinct segments matched in duration to the adjectives.

Procedure

During the experiment, subjects performed 4 separate blocks of trials, 2 in each modality and 2 of each task. Blocks of the same task were always presented sequentially and modality order within each task was constant for each subject. Overall order of tasks and modality order within tasks was counterbalanced across all subjects. Thus, there were 4 possible orderings of the 4 block types.

Before the experiment, subjects practiced their first task in their first modality outside of the MEG room. Although subjects were made aware of the existence of a second task and modality at this time, no specific instructions regarding the second task were given before the

completion of the first task. Instructions and practice for the second task were then given following the completion of both blocks of the first task, while subjects remained in the machine. Prior to recording, subjects' head shapes were digitized using a Polhemus Fastrak 3D digitizer (Polhemus, VT, United States of America). The digitized head shape was then used to constrain source localization during analysis by co-registering 5 coils located around the face with respect to the MEG sensors. Additionally, electrodes were attached 1 cm to the right of and 1 cm beneath the middle of the right eye in order to record the vertical and horizontal electrooculogram (EOG) and detect blinks. Both electrodes were referenced to the left mastoid.

MEG data were collected using a using a whole-head 157-channel axial gradiometer system (Kanazawa Institute of Technology, Tokyo, Japan) sampling at 1000 Hz with a low-pass filter at 200 Hz and a notch filter at 60 Hz. All visual stimuli besides the target shapes were presented for 300 ms and were each followed by a 300 ms blank screen. Corresponding initial auditory stimuli (i.e. noise, adjectives, and list nouns) were presented for 600 ms while critical nouns were presented for 700 ms. These durations did not vary with the length of the individual stimuli and were followed in both cases by 300 ms of silence. In both modalities, target shapes appeared at the end of each trial and remained onscreen until the subject 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 of 500 ms and a standard deviation of 100 ms (see Fig. 1). The entire recording session lasted approximately 1 h.

Data Acquisition

MEG data from the 100 ms prior to the onset of each critical noun to 600 ms post onset were segmented out for each trial. Raw data were first cleaned of potential artifacts by rejecting trials for which the subject answered either incorrectly or too slowly (defined as more than 2.5 s after the appearance of the target shape), or for which the maximum amplitude exceeded 3000 fT, or for which the subject blinked within the critical time window, as determined by manual inspection of the EOG recordings. Remaining data were then averaged for each subject for each condition and band-pass filtered between 1 and 40 Hz. For inclusion in further analysis, we required that subjects 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 sensory responses. In the visual modality, we required either the M100 or M170 field pattern (e.g. Tarkiainen et al. 1999; Pylkkänen and Marantz 2003) to be present in the time window of 100–200 ms following the critical stimuli. In the auditory modality, subject data were required to show a prominent bilateral M100 response situated over the primary auditory cortex, preceded by a similar component of opposite polarity (e.g. Poeppel et al. 1996). In order to assess this criterion, preliminary grand average waveforms were constructed for each subject by averaging over all conditions. Three subjects failed to meet this requirement and were excluded from further 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, Munich, Germany). The channel noise covariance matrix for each estimate was based upon the 100 ms prior to 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 2 shells 10% and 30% below a smoothed standard brain surface. Regional sources in MEG can be regarded as sources with 2 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 of the source activities of its 2 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 subjects and conditions. Minimum norm

images were depth weighted as well as spatio-temporally weighted using a signal subspace correlation measure (Mosher and Leahy 1998).

Data Analysis

We determined effects related to basic combinatory processing using a two-stage analysis. Because the timing of apparently functionally equivalent effects can vary between modalities (e.g. Holcomb and Neville 1990; Osterhout and Holcomb 1993; Marinkovic et al. 2003), to identify shared combinatorial neural activity we first performed a full-brain analysis over activity from each cortical source within each modality separately and then collapsed the results over time. This process produced a full-brain map for each modality that identified which sources were preferentially active during combination (i.e. which sources showed significantly more activity during the 2-word condition during the composition task and no difference during the list task), regardless of the time at which the combinatorial activity occurred. We then merged the results of these two analyses into a single map, assigning each cortical source 1 of 4 possible designations: combinatorially important regardless of modality, combinatorially important for reading only, combinatorially important for speech only, or not combinatorially important. Next, we investigated the time-course of these effects by dividing the resulting map into separate regions of interest (ROIs) and performing a cluster-based permutation test (Maris and Oostenveld 2007) within each region for each modality. This test produces clusters of time points for which the activity localized to a given region was significantly greater during composition compared with the non-combinatorial controls. Thus, our two-step analysis allowed us to identify sources of combinatorial activity that were shared between modalities even if their associated effects occurred at different times for each modality. The follow-up ROI analyses then allowed us to determine the temporal profile within each modality of the identified combinatorial effects.

For the initial full-brain comparison, we first performed a paired t-test within each task at every time point during the critical interval for every cortical source. We then classified sources as significantly active during a task if there was a cluster of at least 20 ms for which the p-value of each t-test was less than 0.05 and the 2-word condition had more activity than the 1-word condition. Thus, based upon this criterion, each source was initially classified as either significantly active or not for each of the 4 tasks. Within each modality, a source was then considered only preferentially active during combinatorial processing if it was significantly active in the composition task and not the list task for that modality. Finally, the resulting full-brain maps for each modality were combined to produce a single map that assigned each cortical source a status as (1) preferentially active during combinatorial processing regardless of modality, (2) preferentially active during visual combinatorial processing only, (3) preferentially active during auditory combinatorial processing only, or (4) not preferentially active during combination.

Note that the initial classification of each source—as preferentially activity during combination in either modality—was arrived at without directly comparing activity measures in the two 2-word tasks to each other. In other words, at no point is the *red boat* condition compared directly with the *cup, boat* condition in our analysis; instead each 2-word condition is first contrasted with the paired 1-word control and the resulting differences are then compared in order to designate each source as combinatorial or not. The reason for this approach, as opposed to performing direct comparisons between the two 2-word conditions, follows from the blocked nature of our design. The purpose of the list task is to ascertain if increased activity identified during the 2-word composition condition is actually due to combinatorial processing or merely the processing of two lexical items during a single trial. Thus, it is constructed to be maximally parallel to the composition task in this latter respect. This parallelism, however, comes at the cost of requiring the subjects to perform a different task. Therefore, it is possible that the demands of this alternate task itself might introduce additional activity into the critical time period, thus confounding any direct intertask comparisons, such as between the *red boat* and *cup, boat* conditions. To mitigate this potential confound, we first contrasted pairs of conditions within each task and only afterwards did we compare combinatorial processing with list processing. Under this approach, activity due simply to the processing of multiple lexical items will still emerge in both the composition and list comparisons, as any increases in 2-word over 1-word activity will be carried forward to the next stage of analysis; however, increases in activity due solely to the demands of one task or the other will be neutralized before the critical between-task comparison. Thus, while the list task serves as a control for the composition task, it is the task as a whole that serves as the control and not the 2-word list condition by itself.

In the second analysis stage, we constructed ROIs from the resulting full-brain analysis by extracting all clusters of combinatorial sources that spanned at least 10 adjacent cortical locations. The time-course of activity localized to each region was then investigated within each modality using the same cluster-based permutation test as in our previous work (see Bemis and Pylkkänen 2011 for details). Briefly, this test analyzes the interaction of activity from the 4 conditions in order to identify time intervals for which there was significant combinatorial activity in a given region, corrected for multiple comparisons across the entire epoch. Additionally, we performed simpler follow-up cluster tests within each task to support these analyses (see Bemis and Pylkkänen 2011 for details). As the ROIs had already been identified as combinatorial during the first phase of the analysis, the primary purpose of these ROI analyses was to characterize and compare the timing and strength of the combinatorial effects within each modality.

Results

Behavioral

Reaction time and accuracy data were subjected to a $2 \times 2 \times 2$ repeated-measures ANOVA with Modality (Visual vs. Auditory), Task (Composition vs. List), and Words (1 vs. 2) as factors (see Fig. 2). The only significant effect for accuracy was a main effect of modality ($F_{1,18} = 9.394$, $P = 0.007$). Subjects were more accurate in the auditory modality compared with the visual modality (97.7% avg. [1.93% std.] vs. 95.6% avg. [6.42% std.]). Although no other effects were significant, it should be noted that accuracy data were qualitatively different between modalities. Within the Composition task, accuracy did not differ significantly within the visual modality though responses tended to be more accurate for the 2-word condition than the 1-word condition ($P = 0.28$; 97.4% avg. [3.3% std.] vs. 94.3% avg. [10.8% std.]). For the auditory modality, on the other hand, subjects were significantly *less* accurate for the 2-word Composition trials compared with the 1-word condition ($P = 0.004$; 97.2% avg. [2.0% std.] vs. 98.7% avg. [1.3% std.]). Within the List task, accuracy results were virtually identical across modalities. Subjects were more accurate for 1-word trials in both modalities (Visual: $P = 0.052$, 94.0% avg. [5.5% std.] vs. 96.7% avg. [2.2% std.]; Auditory: $P = 0.002$, 96.3% avg. [2.0% std.] vs. 98.5% avg. [1.4% std.]). The lack of significance for these effects within the larger ANOVA is likely due to the subjects being close to ceiling on all tasks.

Unlike accuracy, reaction times were qualitatively similar across modalities and exhibited a significant interaction between Task and Number of words ($F_{1,18} = 42.328$, $P < 0.001$). This interaction was driven by a significant decrease in response times for 2-word Composition trials relative to 1-word Composition trials ($P = 0.012$; 602 ms avg. [121 ms std.] vs. 634 ms avg. [112 ms std.]) and a significant increase in response times for 2-word List trials relative to 1-word List trials ($P < 0.0001$; 728 ms avg. [152 ms std.] vs. 650 ms avg.
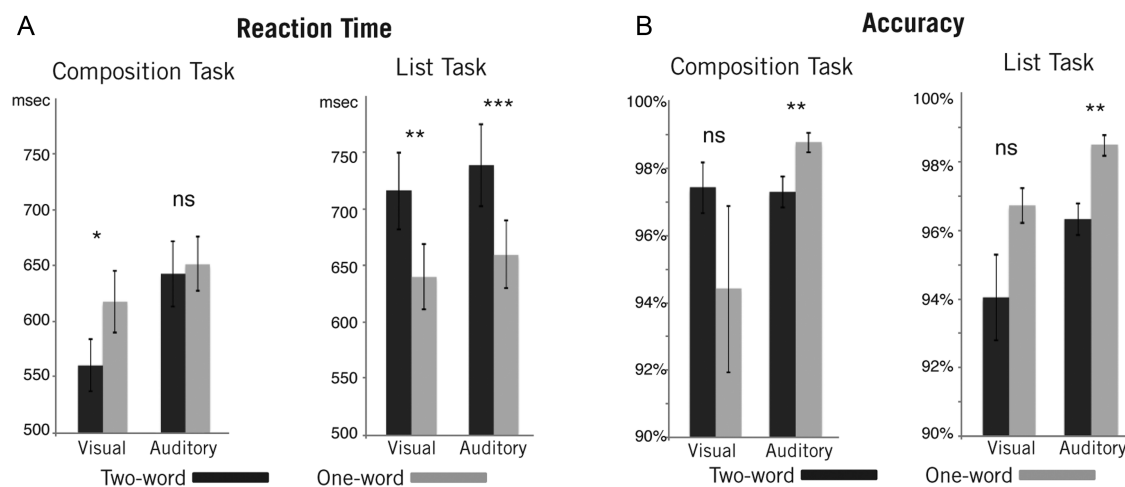
Figure 2. Behavioral results. (*A*) Reaction time and (*B*) accuracy data were submitted to a 2 × 2 × 2 repeated-measures ANOVA with modality (visual vs. auditory), task (composition vs. list), and number of words (1 vs. 2) as factors. We observed a significant interaction between the task and number of words for reaction time ($F_{1,18} = 42.33$, $P < 0.001$), with post hoc tests revealing slower responses in the 2-word list conditions and faster responses in the 2-word composition conditions compared with matched 1-word controls. For accuracy, there was a main effect of modality, with subjects more accurate in the auditory modality compared with the visual modality ($F_{1,18} = 9.40$, $P = 0.007$). Overall, the results indicate that 2-word trials were harder than 1-word trials for every task except visual composition, for which the opposite result held. ns, Nonsignificant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

[126 ms std.]). Within the Composition task, the increase in reaction time for 1-word trials was significant in the visual modality ($P = 0.010$; 561 ms avg. [102 ms std.] vs. 617 ms avg. [120 ms std.]) but not the auditory modality ($P = 0.52$; 643 ms avg. [127 ms std.] vs. 651 ms avg. [105 ms std.]). In both modalities, reaction times were significantly slower for 2-word List responses compared with 1-word List decisions (Visual: $P = 0.003$; 717 ms avg. [149 ms std.] vs. 641 ms avg. [126 ms std.]; Auditory: $P < 0.001$; 739 ms avg. [158 ms std.] vs. 659 ms avg. [129 ms std.]).

Overall, behavioral results in the visual modality replicated those found in our previous experiment (Bemis and Pylkkänen 2011). In the visual List task, 2-word trials were harder than 1-word trials, producing both slower and less accurate responses. As before, this pattern was reversed within the Composition task, with 2-word trials proving easier than 1-word trials. In the auditory modality, results were essentially the same for the List task with 2-word decisions coming both significantly slower and less accurately than 1-word responses, however, within the Composition task the auditory responses showed no clear difference in difficulty between the two conditions. Two-word trials were responded to slightly faster than 1-word trials, but subjects were significantly less accurate in doing so. The cause of this difference across modalities within the Composition task is unclear. One possibility is that this discrepancy reflects a general slowdown in auditory processing relative to visual processing (e.g. Marinkovic et al. 2003) such that the advantage of combination seen for the visual modality might not be available as readily within the auditory modality.

General Assessment of MEG Data

Figure 3 depicts a qualitative overview of the MEG responses evoked during the critical items in each modality, averaged across all subjects. In the visual modality, canonical responses can be seen at ~100 and ~150 ms following the presentation of the critical word. These responses have been identified following the presentation of most complex visual stimuli and

are thought to reflect early sensory and secondary processing of the visual word form (Tarkiainen et al. 1999; Pylkkänen and Marantz 2003). As can be seen, cortical activity during this time is localized to the primary visual cortex and inferior temporal-parietal cortex. These responses are then followed by the characteristic M250 and M350 patterns, which are consistently evoked by the presentation of visual words (Embick et al. 2001; Pylkkänen et al. 2002). As expected (Marinkovic et al. 2003), activity during this later time window localizes primarily to the left temporal cortex.

In the auditory modality, characteristic early field patterns can be observed at ~50 and ~100 ms following the onset of the spoken word. As in previous work (e.g. Poeppel et al. 1996), we see that these early auditory responses are very similar in spatial extent and location but have opposite polarities. Activity during this time period localizes to the primary auditory cortex. As in the visual modality, activity then spreads across the temporal lobe and localizes to similar but slightly larger cortical regions. Thus, across both modalities the general pattern of activity very closely mirrors that observed in previous MEG comparisons of visual and auditory word processing (e.g. Marinkovic et al. 2003). Evoked activity initially localizes to the corresponding primary and associative cortices for each modality and then converges, largely within the temporal lobe, following these early responses.

Full-Brain Combinatorial Analysis

The results of our full-brain analysis are shown in Figure 4A. We identified two cortical regions that reflect combinatorial activity in both modalities—one in the left anterior temporal lobe (lATL) and one in the left angular gyrus (lAG). Additionally, we observed a large region in the right posterior temporal and occipital cortices that reflected combinatorial activity in the visual modality alone. Auditory specific combinatorial sources were located anterior to the shared lAG region, near the primary auditory cortex. Additionally, a separate auditory-specific source abutting the inferior edge of the lAG region was identified. For simplicity, this latter source
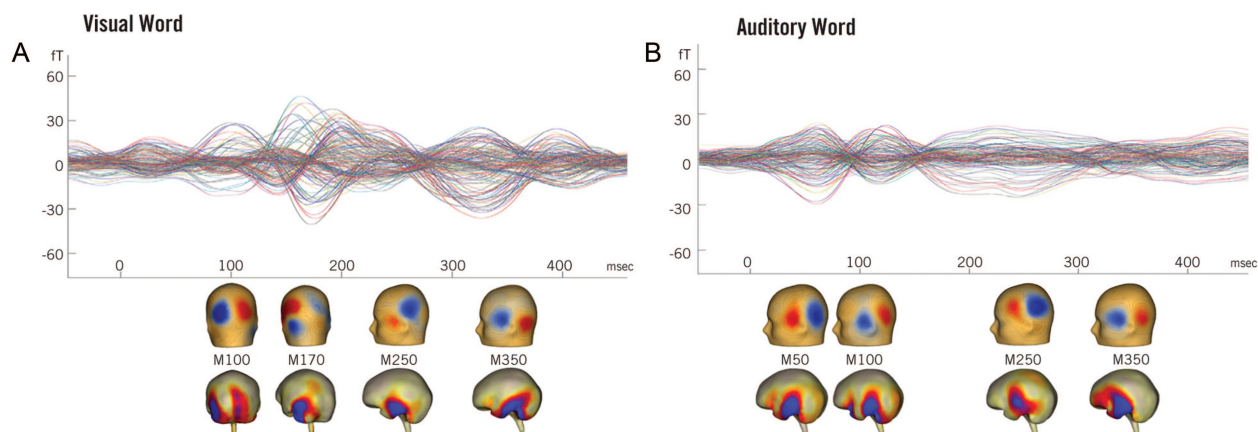
Figure 3. MEG sensor data. The average evoked response to the critical noun is displayed for each modality, collapsed across conditions and subjects. (*A*) Following the presentation of a visual word, canonical evoked response peaks and field patterns (M100 and M170) (Tarkiainen et al. 1999; Pylkkänen and Marantz 2003) are visible at ~100 and ~150 ms, localized to the primary visual cortex and fusiform gyrus respectively. These components are followed by the M250 and M350 peaks and field patterns, which localize to the left temporal lobe and are consistently observed during language comprehension (Embick et al. 2001; Pylkkänen et al. 2002). (*B*) Following the presentation of an auditory word, distinct M50 and M100 field patterns (Poeppel et al. 1996) can be seen, localized to the primary auditory cortex. These are also followed by the M250 and M350 components, again localized to the left temporal lobe.

was incorporated into the larger lAG ROI in the subsequent temporal analysis.

Left Anterior Temporal Lobe ROI

Our interaction cluster test identified a significant time region of combinatorial activity within the lATL (see Fig. 4*B*) for the visual modality from 193 to 241 ms ($P = 0.018$) and in the auditory modality from 267 to 357 ms ($P = 0.049$). Follow-up permutation tests within each task supported these results and identified a significant cluster of combinatorial activity in the visual composition task (191–299 ms, $P = 0.006$) and a marginally significant cluster of combinatorial activity within the auditory composition task (268–323 ms, $P = 0.086$). Neither test within the list task revealed any time windows for which activity during the 2-word condition was substantially greater than the 1-word condition (visual: all clusters with $P > 0.25$; auditory: all clusters with $P > 0.8$).
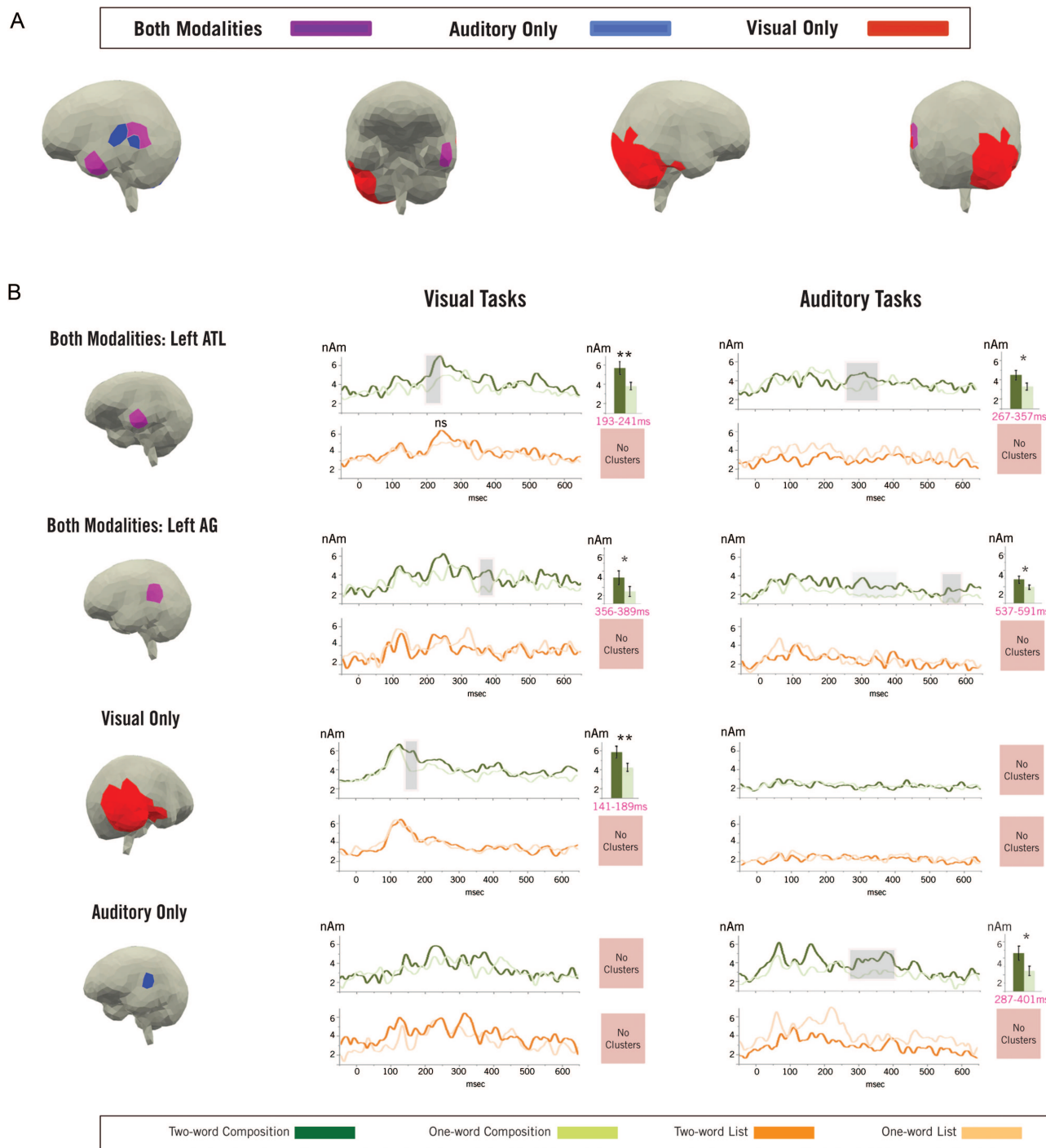
A visual inspection of activity in the lATL, however, does appear to show an increase in 2-word activity during the visual list task in a similar time window as the combinatorial effect identified above. Although this effect does not approach significance in any of our multiple-comparison tests, a post hoc analysis directed specifically at this task and time interval does show a marginally significant increase in activity during the 2-word list condition compared with the 1-word control (1-tailed *t*-test, 200–300 ms, $P = 0.06$). This increase might reflect the fact that, along with combinatorial processing, the lATL has also been linked to the semantic processing of single words (e.g. Devlin et al. 2000; Rogers et al. 2006; Visser et al. 2010). Thus, it may be the case that this cortical region performs multiple functions during the comprehension of language, both associated with and orthogonal to combinatorial processing, and that the interaction of these multiple processes results in a combinatorial effect being overlaid upon more muted lexical-semantic process. Future research will hopefully be able to disentangle these potentially different processes and move towards a reconciliation of these, as of now, disparate literatures.

Overall, our analysis of activity localized to the lATL reveals that combinatorial processing preferentially drove activity

within this region in both modalities. This effect was both earlier and stronger during reading compared with listening. An inspection of the waveforms shows that the auditory lATL combinatorial activity peaks approximately 100 ms after the visual activity and is more temporally extended than its visual counterpart. Qualitatively, these results are similar to many past results that find auditory effects to both begin later and last longer than their visual counterparts (e.g. Marinkovic et al. 2003).

Angular Gyrus ROI

Within the lAG ROI (see Fig. 4*B*), the interaction test revealed a cluster of combinatorial activity in the visual tasks that was marginally significant from 356 to 389 ms ($P = 0.074$). A follow-up test within the composition task alone confirmed this effect, showing a similar cluster of combinatorial activity from 336 to 390 ms ($P = 0.082$). No such cluster was observed in the list task (all clusters $P > 0.65$). In the auditory modality, a significant cluster of combinatorial activity was identified from 537 to 591 ms ($P = 0.040$). (Additionally, a marginally significant cluster was identified from 283 to 326 ms, $P = 0.052$; however, this activity was contiguous both in time and space with an effect found in the auditory specific sources, and so it is discussed below in that context.) The follow-up test within the composition task alone did not reach significance due to a 9 ms gap in the significant cluster identified by the interaction test for which the pairwise comparisons fell above the established cluster threshold. However, a targeted comparison within the composition task at the time identified by the interaction test did show significantly greater activity in the 2-word auditory composition condition compared with the 1-word control ($P = 0.05$, paired *t*-test). Within the list task, no clusters of increased 2-word activity were identified by a follow-up cluster test (all clusters $P > 0.65$), and no significant increases were found for any 50 ms time window between 400 and 600 ms (tested at 10 ms intervals, all tests $P > 0.20$). Thus, our results indicate the involvement of the lAG in combinatorial processing following lATL activity in both modalities. In the visual modality, combinatorial lAG activity followed lATL activity by approximately 150 ms, while

Figure 4. Source analysis. (*A*) A full-brain map indicating the sources that were found to be preferentially active during combination for the different modalities. (*B*) The results of targeted ROI analyses on the activity localized to each ROI identified in the full-brain analysis. Rows represent activity from each ROI, columns represent activity from each modality, and each pair of graphs represents activity from the Composition (green lines) and List (orange lines) tasks for a single ROI and modality. Clusters of combinatorial activity are shaded in gray. In the left ATL, we found significant combinatorial activity in both the visual (193–241 ms; $P = 0.018$) and auditory (267–357 ms; $P = 0.049$) modalities. In the left AG, we found significant combinatorial activity in the auditory modality (537–591 ms; $P = 0.040$) and marginally significant combinatorial activity in the visual modality (356–389 ms; $P = 0.074$). (The light gray shading indicates a marginal combinatorial effect discussed further in the *Results* section). Activity in a large portion of the right occipital-temporal cortex showed significant combinatorial activity in the visual modality (141–189 ms; $P = 0.011$) and not the auditory modality, while activity near the primary auditory cortex displayed the opposite pattern of results (significant auditory cluster: 287–401 ms; $P = 0.042$). No comparison within any List task showed any significant effect. ns, Nonsignificant; $*P < 0.05$; $**P < 0.01$.

in the auditory modality the delay was slightly longer, ~250 ms. As with the lATL, this general activity profile, with auditory effects both slower and more elongated than visual effects, matches previous findings (e.g. Marinkovic et al. 2003).

Visual Specific Combinatorial ROI

Our initial full-brain analysis identified a large region of activity in the right occipital-temporal cortex that was preferentially active during the visual composition task. A cluster

analysis within this region (see Fig. 4*B*) revealed significant combinatorial activity in the visual modality from 141 to 189 ms (*P* = 0.011). A similar cluster was also identified within the visual composition task alone, 134–202 ms (*P* = 0.022). No such cluster was identified within the visual list task (all clusters with *P* > 0.80). Within the auditory modality, no time periods showed any type of effect in this region for any of the cluster tests (all clusters *P* > 0.35). While the observed effect certainly extends across a relatively large section of cortex, the perceptual analysis of visual features has been shown to drive cortical activity across a similar spatial extent in both fMRI (e.g. Grill-Spector et al. 2001) and MEG (e.g. Solomyak and Marantz 2010). Thus, our results indicate that combinatorial effects in this region occurred approximately in the location and time window most often associated with the perceptual parsing of complex visual objects (e.g. Tarkiainen et al. 1999; Solomyak and Marantz 2009) and were restricted to the visual modality.

Auditory Specific Combinatorial ROI

The region identified in the full-brain analysis as active only during auditory combination fell close to the primary auditory cortex and was adjacent to the shared lAG ROI (see Fig. 4*B*). A cluster analysis of this region within the auditory modality revealed a significant cluster of combinatorial activity from 287 to 401 ms (*P* = 0.039) that was again confirmed by a follow-up test within the composition task alone (286–419 ms, *P* = 0.042). No clusters of increased 2-word activity were identified in the auditory list task. Within the visual modality, no significant clusters of combinatorial activity were identified by either the interaction or modality specific permutation tests (all clusters *P* > 0.15). As noted above, the auditory effect identified in this region temporally coincides with a marginally significant combinatorial lAG effect. We believe that this latter activity is an extension of the effect identified in this auditory-specific ROI as it occurs in sources adjacent to the auditory-specific ROI but is shorter in time (114 ms vs. 43 ms) and smaller in strength (*P* = 0.04 vs. *P* = 0.05). It is possible that lAG activity during this time reflects a separate combinatorial mechanism that happens to be active at the same time in nearly the same place, however, the spatial detail of MEG is not sufficient to fully decide between these two possibilities. We believe that the present data are most consistent with one combinatorial effect in the auditory modality from approximately 300 to 400 ms, centered near the primary auditory cortex but potentially spreading into the lAG ROI, and another later combinatorial effect from approximately 550 to 600 ms, centered in the lAG.

Discussion

The present study constitutes the first direct comparison of the neural correlates of basic linguistic combinatorial processing in both reading and listening. While there have been previous investigations into shared combinatorial mechanisms in language comprehension (e.g. Michael et al. 2001; Constable et al. 2004), none have reduced the scope of the studied processes to that of a simple adjective–noun phrase. In the present paradigm, we used MEG to record neural activity as subjects either listened to or read common, object-denoting nouns (*boat*) and measured increases in this activity during minimal compositional contexts (*red boat*) compared with

non-compositional controls (*xhl boat*). To account for effects due solely to differences in lexical-semantic material during the combinatorial condition, we also compared a non-combinatorial list condition (*cup, boat*) with the same non-compositional control. In both modalities, we identified increased activity in the left anterior temporal lobe (lATL) and left angular gyrus (lAG) associated with basic linguistic combination. A subsequent targeted ROI analysis within each modality revealed a temporal consistency to these effects such that combinatorial lATL activity preceded combinatorial lAG activity in both modalities. Between modalities, differential effects were confined to primary and associative sensory cortical regions and occurred prior to modality-independent combinatorial activity.

Thus, our results suggest that both auditory and visual processing of minimal linguistic phrases relies upon a shared set of basic combinatorial neural mechanisms that localize to the lATL and lAG and that shared combinatorial effects previously identified in these regions during more complex processing (e.g. Humphries et al. 2001; Vandenberghe et al. 2002) reflect basic combinatorial mechanisms. Additionally, early perceptual processing appears to be modulated during the comprehension of basic linguistic phrases in both modalities.

Modality-Independent Combinatorial Effects

Our initial full-brain analysis identified two neural regions—the lATL and lAG—that were significantly more active during composition compared with non-compositional controls in both the visual and auditory modalities. Targeted ROI analyses within both regions revealed clusters of time points for which activity elicited by the critical noun was significantly greater during composition than the non-combinatorial control. This result held for both modalities, and in both modalities increased lATL activity preceded increased lAG activity during basic combinatorial processing. Importantly, no significant increases were identified for activity within these regions during the non-compositional contrast at any point in either modality. In the visual modality, the increased lATL activity peaked at approximately 225 ms (with a significant cluster of combinatorial activity occurring from 193 to 241 ms) followed by increased lAG activity at approximately 375 ms (356–389 ms cluster). In the auditory modality, the combinatorial lATL effect occurred at approximately 300 ms (267–357 ms cluster) followed by increased lAG activity at ~560 ms (537–591 ms cluster). One likely explanation for the delay in auditory processing is that during listening a comprehender must wait for a word to unfold in time before it can be uniquely identified relative to its phonological competitors whereas in reading the entire word is present immediately. Typical models of auditory word recognition posit that higher-level processing must wait until this uniqueness point has been reached (e.g. Marslen-Wilson 1987), and the temporal delay associated with waiting for the uniqueness point has been shown to affect the timing of electrophysiological components such as the N400 (O'Rourke and Holcomb 2002). Thus, the temporal discrepancy we observe here might reflect the extra time needed during listening to accrue enough information from the auditory signal to uniquely identify the incoming word—a suggestion offered previously to explain delays observed between auditory effects and their visual counterparts (e.g. Marinkovic et al. 2003; Leinonen et al.

2009). Thus, our results suggest that visual and auditory language processing utilize shared combinatorial mechanisms during the comprehension of simple linguistic phrases and that these shared operations are subserved by the lATL and lAG.

Previous work has implicated both of these regions in combinatorial linguistic processing. The lATL has consistently exhibited more activity during the processing of sentences compared with word lists (Mazoyer et al. 1993; Bottini et al. 1994; Stowe et al. 1998; Friederici et al. 2000; Humphries et al. 2001; Vandenberghe et al. 2002) and was found to be significantly involved during basic composition in a previous, visual-only study using the present paradigm (Bemis and Pylkkänen 2011). Thus, the current results bolster the hypothesis that effects found in the lATL during sentential processing reflect basic combinatorial operations. Furthermore, they suggest that the neural mechanisms underlying these processes are shared between the auditory and visual modalities; a finding consistent with previous results derived from more complex expressions (e.g. Jobard et al. 2007; Lindenberg and Scheef 2007).

The lAG has also been implicated during the comprehension of sentences compared with word lists (Bottini et al. 1994; Humphries et al. 2001; Humphries et al. 2005; Humphries et al. 2007), and, as with the lATL, was found to be preferentially active during basic composition in our previous study, though this effect was relatively weak (P = 0.11, Bemis and Pylkkänen 2011). Additionally, neurological studies indicate that lesions in the lAG are associated with difficulties in processing complex sentences (Dronkers et al. 2004), leading to the suggestion that this region may play a role in integrating semantic information into context (Lau et al. 2008). Thus, the present results suggest that previous sentential-level effects observed in the lAG also reflect basic combinatorial operations and that these processes operate independent of the modality of input.

Beyond the shared spatial nature of these two effects, the consistency in their temporal ordering—lATL activity preceded lAG activity in both modalities—further supports the conclusion that these effects reflect shared combinatorial operations. While the gross nature of our contrast makes it difficult to parcel out different types of combinatorial operations to one region or the other, many existing models of sentence comprehension posit a two-step process in which syntactic composition precedes semantic combinatorial operations (e.g. Friederici 2002). The lATL has previously been linked to syntactic processing, with increased activity in this region corresponding to measures of syntactic complexity during natural story comprehension (Brennan et al. 2012) and decreased activity resulting from contexts that elicit syntactic priming (Noppeney and Price 2004). Contrastingly, the lAG has been found to be significantly more active during the comprehension of semantically congruent sentences compared with syntactically well formed but incoherent sentences (Humphries et al. 2007), and has thus been suggested to be critically involved in the integration of semantic information during language processing (Lau et al. 2008). While the present results are consistent with a delineation of syntactic combinatorial processing to the lATL and semantic composition to the lAG, this suggestion must remain extremely tentative at the present time. Several recent studies have also associated activity in at least certain regions of the lATL with

combinatorial semantic operations (Rogalsky and Hickok 2009; Baron et al. 2010; Baron and Osherson 2011; Pallier et al. 2011), while increased lAG activity has been linked to syntactic processing (Dapretto and Bookheimer 1999). Thus, as the present manipulation does not allow for direct insight into the partitioning of distinct combinatorial operations to distinct cortical regions, more definitive conclusions must await more directed investigations into this specific question.

Finally, it should be noted that while our previous study (Bemis and Pylkkänen 2011) also identified a two-phase pattern of combinatorial activity in which the first phase consisted of increased lATL activity, the location of the most salient subsequent combinatorial effect was in the ventromedial prefrontal cortex (vmPFC) and not the lAG. A targeted post hoc analysis of vmPFC activity in the present results, based on the spatial-temporal ROI identified in the previous study, did reveal a significant increase in activity during the 2-word condition compared with the 1-word control within the composition task (paired t-test, 300–400 ms, $P < 0.05$). However, in the present study, this increase was accompanied by a corresponding increase in the list task as well (paired t-test, 300–400 ms, $P < 0.01$). Thus, the strong interaction observed in the previous study was not replicated in the present results. Within the auditory domain, vmPFC activity exhibited the pattern predicted for combinatorial activity (i.e. activity in the 2-word composition condition exceeded that in the 1-word control to a greater degree than in the list task), however, results from a targeted post hoc analysis of this activity did not reach significance (300–400 ms, repeated-measures ANOVA, all $F < 2$; paired t-tests within tasks, all $t < 1$). Thus, although our present results are not inconsistent with our past findings, they less clearly indicate combinatorial processing in the vmPFC, as the significant increase observed in vmPFC activity during the composition task was also accompanied by a corresponding increase in the list task as well.

There are several possible explanations for this unexpected result. For one, the functional theory underlying the list task is underspecified relative to the composition task, and so unexpected increases in activity might occur during the former depending upon how subjects perform this task. Perhaps a more completely articulated functional theory of the list task in conjunction with an analysis of different behavioral metrics might help in explaining the differences observed in vmPFC activity. At the present time, however, the exact form of such an explanation remains unclear. Another possibility is that the increased length of the experiment (nearly twice as long as the original) as well as the interleaving of multiple modalities and tasks might have contributed to modulations of activity within the vmPFC, as task switching has been shown to affect activity in this region (e.g. Dreher et al. 2002) as well as in nearby cortical and subcortical structures (e.g. Swainson et al. 2003; Crone et al. 2006). Unfortunately, a post hoc analysis breaking down different task and modality orders lacked sufficient power to identify any significant effects that might support this hypothesis. Another deviation from the previous experiment was that in the present analysis we used a more conservative artifact rejection protocol, both in terms of a stricter amplitude rejection criterion and the use of ocular EOG measurements to reject trials containing blinks. However, a re-analysis of our present data using exactly the same parameters as in the previous study did not change the qualitative pattern

of the results. Finally, MEG activity localized to the vmPFC is susceptible to several sources of variability that do not affect other cortical regions (such as the lATL and lAG) to nearly the same degree, and these difficulties might create greater instability in vmPFC effects. For example, it has been demonstrated that eye movements can contaminate MEG measurements of neural activity from the vmPFC and surrounding regions (Carl et al. 2012). Although our results were not affected by the inclusion or exclusion of trials with excessive EOG amplitudes, subtle eye movements that fall beyond the ability of EOG to detect can appear similar to neural activity in both visual (Yuval-Greenberg et al. 2008) and auditory (Yuval-Greenberg and Deouell 2011) electrophysiological paradigms. It should be noted, however, that these distortions are not believed to be time-locked to the stimulus and have previously manifested themselves as induced rather than evoked effects. Lastly, as pointed out in our previous paper, the ability of MEG to reliably detect and localize cortical activity decreases with the distance from the sensors (Hillebrand and Barnes 2002), which may contribute to more variable measurements from regions such as the vmPFC compared with locations closer to the surface, such as the lATL and lAG. So, in sum, there are several potential explanations for why the present results did not clearly replicate the combinatorial effect in the vmPFC that was previous observed, however, at the present time none of these theories unambiguously supply a complete explanation for this discrepancy. Clearly, more work is needed in order to disentangle how different neural generators contribute to later combinatorial effects.

On the other hand, there is now a growing body of converging evidence that implicates the lATL in basic linguistic combination. The present results further strengthen this conclusion and indicate that the cognitive operation subserved by this neural mechanism is utilized by both auditory and visual language processing.

Modality-Specific Combinatorial Effects

For both modalities, we observed significant modality-specific combinatorial activity that localized close to the primary cortices of the associated modality. During reading, we observed a large region of combinatorial activity located in the right occipital-temporal cortex. The location and timing (approximately 160 ms following the onset of the noun) of this effect corresponds roughly to the secondary sensory response that is thought to occur following all visually complex stimuli—the M170 response (e.g. Tarkiainen et al. 1999; Solomyak and Marantz 2010). This component is hypothesized to reflect the analysis of orthographic and morphological properties of a visual word form, as it responds preferentially to strings of letters compared with symbols (Tarkiainen et al. 1999), correlates with properties of the visual form of a word (Solomyak and Marantz 2010), and reflects early morphological decomposition processes (Zweig and Pylkkänen 2009). Thus, our results suggest that in the present study basic combination in the visual modality was accompanied by increased processing related to the analysis of the visual form of the combined noun.

During listening, modality-specific combinatorial sources were also identified close to the primary sensory cortex, this time centered over the left auditory cortex. Both the location, adjacent to the lAG ROI, and the timing of this effect,

concurrent with a marginally significant combinatorial effect in the lAG ROI at ~300 ms, suggest that this effect reflects modality-specific combinatoric activity that occurs across a rather large cortical region encompassing both the primary auditory cortex as well as part or all of the lAG ROI, and that precedes subsequent combinatorial activity in the latter region. A similar pattern of results was observed during a cross-modal comparison of anomaly processing (Hagoort and Brown 2000b) in which the N400 effect, observed in both modalities, was preceded by a topographically similar auditory-specific combinatorial effect. Taken together, these findings suggest that previous results reporting an earlier onset for auditory compared with visual combinatorial processing (e.g. Holcomb and Neville 1991; Osterhout and Holcomb 1993) might in fact reflect two separate auditory mechanisms—the first specific to auditory processing and the second shared between modalities (cf. Hagoort and Brown 2000b). As previous MEG studies have generally localized processing of the acoustic form of an auditory word to a cortical region similar to the present result (see Salmelin 2007) our results suggest that this first, auditory specific effect is associated with initial perceptual processing and that, as in visual processing, auditory combination in the current study was also accompanied by increased processing of the critical noun's sensory form—in this case, its acoustic form.

Thus, at a general level, our modality-specific findings were both relatively unsurprising and consistent across modalities. For both reading and listening, we found early combinatorial effects specific to each modality that were located near the associated primary and secondary sensory cortices and were reminiscent of past components thought to reflect the analysis of a word's sensory form. The importance of top-down, predictive information during the processing of combinatorial linguistic phrases has long been emphasized both theoretically (e.g. McClelland 1987; Hale 2003; Levy 2008) and empirically (e.g. Trueswell 1996; Spivey and Tanenhaus 1998; Gibson 2006). Additionally, top-down informational flow has been shown to modulate neural responses generated by the perceptual processing of both auditory (e.g. Debener et al. 2003) and visual (e.g. Engel et al. 2001) stimuli. Recently, MEG work investigating the interplay of these factors has shown that the early perceptual responses to a word can be modulated by the top-down prediction of its grammatical category (Dikker et al. 2009; Dikker et al. 2010). The results of the present study support this general framework and suggest that past differences observed between auditory and visual combinatorial effects might in part reflect top-down modulations of perceptual processing that appear to be active even during the comprehension of minimal linguistic phrases. Specifically, however, the functional significance of these modality-specific effects is unclear. It may be the case that the expectation of combination or the engagement of a grammatical parser directs more resources toward the initial processing of an impending word. However, at the present time, hypotheses regarding the specific mechanisms underlying these effects must remain conjecture until more targeted investigations directed specifically at this phenomenon have been carried out.

Although we did not identify any additional modality-specific effects of combination, it should be noted that this might be due in part to the sequential presentation of visual stimuli used in the present paradigm. While presenting

written words one at a time allowed us to precisely isolate activity evoked by the critical noun and to control the critical stimuli across conditions, it also created a similarity between the two modalities—namely, the sequential presentation of individual lexical items—that does not necessarily exist during more natural comprehension, where written expressions are normally encountered as static wholes. While recent evidence indicates that canonical electrophysiological effects persist between these two presentation paradigms (Dimigen et al. 2011), one interesting avenue for future research would be to determine whether a more natural reading paradigm might lead to more pronounced modality-specific combinatorial effects.

Conclusion

Using a linguistically minimal paradigm, we were able to isolate combinatorial neural activity evoked by the construction of simple adjective–noun phrases in both speech and reading. Further, we were able to do so using the same task and linguistic items across modalities. Our results indicate that linguistic processing in both modalities share common combinatorial mechanisms when constructing and understanding basic linguistic phrases. During both reading and listening, we found increased activity during basic composition in the lATL and the lAG. Further, in both cases the combinatorial lATL effect preceded combinatorial activity in the lAG. These results suggest that auditory and visual language processing draw upon the same neural mechanisms during basic combinatorial processing and that these mechanisms are deployed in a similar manner for both. Additionally, we identified modality-specific increases in combinatorial activity that temporally and spatially coincided with early perceptual responses, suggesting a modulation of early perceptual processing within both modalities during the comprehension of minimal linguistic phrases. Further work must now build upon these results in order to determine how closely the correspondence between auditory and visual combinatorial neural mechanisms remains as the complexity of the stimulus increases.

Funding

Notes

References

Anderson JE, Holcomb PJ. 1995. Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. Psychophysiology. 32(2):177–190.

Balconi M, Pozzoli U. 2005. Comprehending semantic and grammatical violations in Italian. N400 and P600 comparison with visual and auditory stimuli. J Psycholinguist Res. 34(1):71–98.

Baron SG, Osherson D. 2011. Evidence for conceptual combination in the left anterior temporal lobe. NeuroImage. 55(4):1847–1852.

Baron SG, Thompson-Schill SL, Weber M, Osherson D. 2010. An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. Cogn Neurosci. 1 (1):44–51.

Beauvois MF. 1982. Optic aphasia: A process of interaction between vision and language. Philos Trans R Soc London B. 298:35–47.

Bemis DK, Pylkkänen L. 2011. Simple composition: An magnetoencephalography investigation into the comprehension of minimal linguistic phrases. J Neurosci. 31(8):2801–2814.

Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM. 2002. Modality independence of word comprehension. Hum Brain Mapp. 16(4):251–261.

Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM. 2003. Relation between brain activation and lexical performance. Hum Brain Mapp. 19(3):155–169.

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.

Brainard DH. 1997. The psychophysics toolbox. Spat Vis. 10 (4):433–436.

Brennan J, Nir Y, Hasson U, Malach R, Heeger DJ, Pylkkänen L. 2012. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 120:163–173.

Buchweitz A, Mason RA, Tomitch LMB, Just MA. 2009. Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. Psychol Neurosci. 2(2):111–123.

Buckner RL, Koutstaal W, Schacter DL, Rosen BR. 2000. Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. Brain. 123:620–640.

Caramazza A, Hillis AE, Rapp BC, Romani C. 1990. The multiple semantics hypothesis: Multiple confusions? Cogn Neuropsychol. 7(3):161–189.

Carl C, Açık A, König P, Engel AK, Hipp JF. 2012. The saccadic spike artifact in MEG. NeuroImage. 59:1657–1667.

Carpentier A, Pugh K, Westerveld M, Studholme C, Skrinjar O, Thompson J, Spencer D, Constable R. 2001. Functional MRI of language processing: Dependence on input modality and temporal lobe epilepsy. Epilepsia. 42(10):1241–1254.

Cohen L, Jobert A, Le Bihan D, Dehaene S. 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. NeuroImage. 23(4):1256–1270.

Constable RT, Pugh KR, Berroya E, Mencl WE, Westerveld M, Ni W, Shankweiler D. 2004. Sentence complexity and input modality effects in sentence comprehension: an fMRI study. NeuroImage. 22:11–21.

Crone EA, Wendelken C, Donohue SE, Bunge SA. 2006. Neural evidence for dissociable components of task-switching. Cereb Cortex. 16:475–486.

Dapretto M, Bookheimer SY. 1999. Form and content: Dissociating Syntax and semantics in sentence comprehension. Neuron. 24 (2):427–432.

Debener S, Herrmann CS, Kranczioch C, Gembris D, Engel AK. 2003. Top-down attentional processing enhances auditory evoked gamma band activity. NeuroReport. 14(5):683–686.

Dehaene S. 2009. Reading in the Brain. New York (NY): Penguin.

Dérouesné J, Beauvois MF. 1979. Phonological processing in reading: data from alexia. J Neurol Neurosurg Psychiatry. 42:1125–1132.

Devauchelle AD, Oppenheim C, Rizzi L, Dehaene S, Pallier C. 2009. Sentence syntax and content in the human temporal lobe: An fMRI adaptation study in auditory and visual modalities. J Cogn Neurosci. 21(5):1000–1012.

Devlin JT, Russell RP, Davis MH, Price CJ, Wilson J, Moss HE, Matthews PM, Tyler LK. 2000. Susceptibility-induced loss of

signal: Comparing PET and fMRI on a semantic task. NeuroImage. 11:589–600.

Dikker S, Rabagliati H, Farmer TA, Pylkkänen L. 2010. Early occipital sensitivity to syntactic category is based on form typicality. Psychol Sci. 21(5):629–634.

Dikker S, Rabagliati H, Pylkkänen L. 2009. Sensitivity to syntax in visual cortex. Cognition. 110:293–321.

Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R. 2011. Coregistration of eye movements and EEG in natural reading: Analyses and review. J Exp Psychol. 140(4):552–572.

Dreher JC, Koechlin E, Ali SO, Grafman J. 2002. The roles of timing and task order during task switching. NeuroImage. 17:95–109.

Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition. 92(1–2):145–177.

Druks J, Shallice T. 2000. Selective preservation of naming from description and the "Restricted Preverbal Message". Brain Lang. 72 (2):100–128.

Embick D, Hackl M, Schaeffer J, Kelepir M, Marantz A. 2001. A magnetoencephalographic component whose latency reflects lexical frequency. Cogn Brain Res. 10(3):345–348.

Engel AK, Fries P, Singer W. 2001. Dynamic predictions: Oscillations and synchrony in top-down processing. Nat Rev Neurosci. 2:704–716.

Friederici AD. 2002. Towards a neural basis of auditory sentence processing. Trends Cogn Sci. 6(2):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(2):289–300.

Gibson E. 2006. The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. J Mem Lang. 54:363–388.

Gomes H, Ritter W, Tartter VC, Vaughan HG, Rosen JJ. 1997. Lexical processing of visually and auditorily presented nouns and verbs: evidence from reaction time and N400 priming data. Cogn Brain Res. 6(2):121–134.

Grill-Spector K, Kourtzi Z, Kanwisher N. 2001. The lateral occipital complex and its role in object recognition. Vision Res. 41:1409–1422.

Hagoort P, Brown CM. 2000a. ERP efects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. Neuropsychologia. 38:1531–1549.

Hagoort P, Brown CM. 2000b. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia. 38(11):1518–1530.

Hale J. 2003. The information conveyed by words in sentences. J Psycholinguist Res. 32(2):101–123.

Hickok G, Poeppel D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition. 92(1–2):67–99.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nat Rev Neurosci. 8(5):393–402.

Hillebrand A, Barnes G. 2002. A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. NeuroImage. 16:638–650.

Holcomb PJ, Coffey SA, Neville HJ. 1992. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. Dev Neuropsychol. 8:203–241.

Holcomb PJ, Neville HJ. 1990. Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. Lang Cogn Processes. 5(4):281–312.

Holcomb PJ, Neville HJ. 1991. Natural speech processing: An analysis using event-related brain potentials. Psychobiology. 19(4):286–300.

Howard D, Patterson K, Wise R, Brown WD, Friston K, Weiller C, Frackowiak R. 1992. The cortical localization of the lexicons. Positron emission tomography evidence. Brain. 115:1769–1782.

Humphries C, Binder JR, Medler DA, Liebenthal E. 2007. Time course of semantic processes during sentence comprehension: an fMRI study. Neuroimage. 36(3):924–932.

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(2):128–138.

Humphries C, Willard K, Buchsbaum B, Hickok G. 2001. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. NeuroReport. 12(8):1749–1752.

Jakimik J, Glenberg A. 1990. Verbal learning meets psycholinguistics: Modality effects in the comprehension of anaphora. J Mem Lang. 29:582–590.

Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N. 2007. Impact of modality and linguistic complexity during reading and listening tasks. NeuroImage. 34(2):784–800.

Kayser J, Fong R, Tenke CE, Bruder GE. 2003. Event-related brain potentials during auditory and visual word recognition memory tasks. Cogn Brain Res. 16:11–25.

Kuriki S, Murase M. 1989. Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. Exp Brain Res. 77(1):127–134.

Kutas M, Hillyard SA. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science. 207(4427):203–205.

Lau EF, Phillips C, Poeppel D. 2008. A cortical network for semantics: (de)constructing the N400. Nat Rev Neurosci. 9(12):920–933.

Lee D, Newman SD. 2010. The effect of presentation paradigm on syntactic processing: An event-related fMRI study. Hum Brain Mapp. 31:65–79.

Leinonen A, Grönholm-Nyman P, Järvenpää M, Söderholm C, Lappi O, Laine M, Krause CM. 2009. Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: Evidence from a morphologically rich language. Brain Res. 1275:54–66.

Levy R. 2008. Expectation-based syntactic comprehension. Cognition. 106:1126–1177.

Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. 2005. Neural substrates of phonemic perception. Cereb Cortex. 15:1621–1631.

Lindenberg R, Scheef L. 2007. Supramodal language comprehension: Role of the left temporal lobe for listening and reading. Neuropsychologia. 45:2407–2415.

Marangolo P, Rinaldi MC, Sabatini U. 2004. Modality-specific naming deficit: Cognitive and neural mechanisms implicated in naming to definition. Neurocase. 10(4):280–289.

Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E. 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. Neuron. 38(3):487–497.

Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods. 164(1):177–190.

Marr D. 1982. Vision. Cambridge (MA): The MIT Press.

Marslen-Wilson WD. 1987. Functional parallelism in spoken word-recognition. In: Frauenfelder UH, Tyler LK. (Ed.), Spoken Word Recognition. Cambridge (MA): The MIT Press. pp. 71–103.

Mazoyer B, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J. 1993. The cortical representation of speech. J Cogn Neurosci. 5(4):467–479.

McClelland JL. 1987. The Case for Interactionism in Language Processing. In: Coltheart M. (Ed.), Attention and Performance XII: The Psychology of Reading. London (United Kingdom): Erlbaum. pp. 3–35.

Michael EB, Keller TA, Carpenter PA, Just MA. 2001. fMRI investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. Hum Brain Mapp. 13:239–252.

Mosher JC, Leahy RM. 1998. Recursive MUSIC: A framework for EEG and MEG source localization. IEEE Trans Biomed Eng. 45 (11):1342–1354.

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature. 385(6615):432–434.

Niznikiewicz M, O'Donnell B, Nestor P, Smith L, Law S, Karapelou M, Shenton M, McCarley R. 1997. ERP assessment of visual and auditory language processing in Schizophrenia. J Abnorm Psychol. 106 (1):85–94.

Noppeney U, Price CJ. 2004. An FMRI study of syntactic adaptation. J Cogn Neurosci. 16(4):702–713.

O'Rourke TB, Holcomb PJ. 2002. Electrophysiological evidence for the efficiency of spoken word processing. Biol Psychol. 60:121–150.

Osterhout L, Holcomb PJ. 1993. Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. Lang Cogn Processes. 8(4):413–437.

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(4):786–803.

Pallier C, Devauchelle AD, Dehaene S. 2011. Cortical representation of the constituent structure of sentences. Proc Natl Acad Sci USA. 108(6):2522–2527.

Parviainen T, Helenius P, Salmelin R. 2005. Cortical differentiation of speech and nonspeech sounds at 100 ms: implications for dyslexia. Cereb Cortex. 15:1054–1063.

Pelli DG. 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis. 10(4):437–442.

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature. 331(6157):585–589.

Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, Poeppel D, McGinnis M, Roberts T. 2000. Auditory cortex accesses phonological categories: An MEG mismatch study. J Cogn Neurosci. 12(6):1038–1055.

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun. 41:245–255.

Poeppel D, Yellin E, Phillips C, Roberts TPL, Rowley HA, Wexler K, Marantz A. 1996. Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. Cogn Brain Res. 4(4):231–242.

Price CJ. 2000. The anatomy of language: Contributions from functional neuroimaging. J Anat. 197(03):335–359.

Price CJ, Winterburn D, Giraud AL, Moore CJ, Noppeney U. 2003. Cortical localisation of the visual and auditory word form areas: A reconsideration of the evidence. Brain Lang. 86(2):272–286.

Pylkkänen L, Marantz A. 2003. Tracking the time course of word recognition with MEG. Trends Cogn Sci. 7(5):187–189.

Pylkkänen 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.

Rogalsky C, Hickok G. 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb Cortex. 19(4):786–796.

Rogers TT, Hocking J, Noppeney U, Mechelli A, Gorno-Tempini ML, Patterson K, Price CJ. 2006. Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. Cogn Affect Behav Neurosci. 6(3):201–213.

Rugg MD, Nieto-Vegas M. 1999. Modality-specific effects of immediate word repetition: electrophysiological evidence. NeuroReport. 10(12):2661–2664.

Saffran EM, Marin OSM. 1977. Reading without phonology: Evidence from aphasia. Q J Exp Psychol. 29(3):515–525.

Salmelin R. 2007. Clinical neurophysiology of language: The MEG approach. Clin Neurophysiol. 118(2):237–254.

Sass K, Krach S, Sachs O, Kircher T. 2009. Lion—tiger—stripes: Neural correlates of indirect semantic priming across processing modalities. NeuroImage. 45(1):224–236.

Shallice T. 1987. Impairments of semantic processing: Multiple dissociations. In: Coltheart M, Sartori G, Job R. (Eds.), The cognitive neuropychology of language. Hillsdale (NJ): Lawrence Erlbaum Associates Inc. pp. 111–127.

Shallice T. 1988. Specialisation within the semantic system. Cogn Neuropsychol. 5(1):133–142.

Sheridan J, Humphreys GW. 1993. A verbal-semantic category-specific recognition impairment. Cogn Neuropsychol. 10(2):143–184.

Solomyak O, Marantz A. 2010. Evidence for early morphological decomposition in visual word recognition. J Cogn Neurosci. 22(9):2042–2057.

Solomyak O, Marantz A. 2009. Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. Brain Lang. 108(3):191–196.

Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS. 2006. Converging language streams in the human temporal lobe. J Neurosci. 26(28):7328–7336.

Spivey MJ, Tanenhaus MK. 1998. Syntactic ambiguity resolution in discourse: modeling the effects of referential context and lexical frequency. J Exp Psychol Learn Mem Cogn. 24(6):1521–1543.

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(13):2995–2999.

Swainson R, Cunnington R, Jackson GM, Rorden C, Peters AM, Morris PG, Jackson SR. 2003. Cognitive control mechanisms revealed by ERP and fMRI: evidence from repeated task-switching. J Cogn Neurosci. 15(6):785–799.

Tarkiainen A, Cornelissen PL, Salmelin R. 2002. Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. Brain. 125:1125–1136.

Tarkiainen A, Helenius P, Hansen PC, Cornelissen PL, Salmelin R. 1999. Dynamics of letter string perception in the human occipito-temporal cortex. Brain. 122:2119–2131.

Trueswell JC. 1996. The role of lexical frequency in syntactic ambiguity resolution. J Mem Lang. 35:566–585.

Vandenberghe R, Nobre AC, Price CJ. 2002. The response of left temporal cortex to sentences. J Cogn Neurosci. 14(4):550–560.

Vartiainen J, Parviainen T, Salmelin R. 2009. Spatiotemporal convergence of semantic processing in reading and speech perception. J Neurosci. 29(29):9271–9280.

Visser M, Jefferies E, Lambon Ralph M. 2010. Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. J Cogn Neurosci. 22(6):1083–1094.

Warrington EK. 1975. The selective impairment of semantic memory. Q J Exp Psychol. 27(4):635–657.

Xu Y, Liu J, Kanwisher N. 2005. The M170 is selective for faces, not for expertise. Neuropsychologia. 43:588–597.

Yuval-Greenberg S, Deouell LY. 2011. Scalp-recorded induced Gamma-Band responses to auditory stimulation and its correlations with saccadic muscle-activity. Brain Topogr. 24:30–39.

Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. 2008. Transient induced Gamma-Band response in EEG as a manifestation of miniature saccades. Neuron. 58:429–441.

Zweig E, Pylkkänen L. 2009. A visual M170 effect of morphological complexity. Lang Cogn Processes. 24(3):412–439.