

Spatiotemporal Signatures of Lexical–Semantic Prediction

Ellen F. Lau^{1,3,4}, Kirsten Weber^{1,4}, Alexandre Gramfort^{2,5}, Matti S. Hämäläinen^{2,6} and Gina R. Kuperberg^{1,4}

¹Department of Psychiatry, ²Department of Radiology, Harvard Medical School and Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA, ³Department of Linguistics, University of Maryland, College Park, MD 20742, USA, ⁴Department of Psychology, Tufts University, Medford, MA 02155, USA, ⁵Department of Signal and Image Processing, Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, Paris, France and ⁶Division of Health Sciences and Technology, Harvard–Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Address correspondence to Dr Ellen Lau, Department of Linguistics, University of Maryland, College Park, MD 20742, USA.
Email: ellenlau@umd.edu

Although there is broad agreement that top-down expectations can facilitate lexical–semantic processing, the mechanisms driving these effects are still unclear. In particular, while previous electroencephalography (EEG) research has demonstrated a reduction in the N400 response to words in a supportive context, it is often challenging to dissociate facilitation due to bottom-up spreading activation from facilitation due to top-down expectations. The goal of the current study was to specifically determine the cortical areas associated with facilitation due to top-down prediction, using magnetoencephalography (MEG) recordings supplemented by EEG and functional magnetic resonance imaging (fMRI) in a semantic priming paradigm. In order to modulate expectation processes while holding context constant, we manipulated the proportion of related pairs across 2 blocks (10 and 50% related). Event-related potential results demonstrated a larger N400 reduction when a related word was predicted, and MEG source localization of activity in this time-window (350–450 ms) localized the differential responses to left anterior temporal cortex. fMRI data from the same participants support the MEG localization, showing contextual facilitation in left anterior superior temporal gyrus for the high expectation block only. Together, these results provide strong evidence that facilitatory effects of lexical–semantic prediction on the electrophysiological response 350–450 ms postonset reflect modulation of activity in left anterior temporal cortex.

Keywords: anterior temporal cortex, expectation, magnetoencephalography, N400, semantic

Introduction

Successful language comprehension requires mapping visual or auditory input to stored lexical and semantic memory representations in order to activate the concepts that the speaker or writer intended to communicate. During typical comprehension, there are several serious potential challenges to accomplishing this task, such as noise and variability in the input and ambiguity in the word-to-concept mapping. A number of authors have proposed that context-based prediction may be part of the solution to these challenges (Wicha et al. 2004; DeLong et al. 2005; Lau et al. 2006; Federmeier 2007; Dikker et al. 2009; Van Petten and Luka 2012). By using context to generate predictions about what input is likely to come next, the processor can more easily overcome these problems and, across time, can allocate resources more efficiently to deal with rapid input.

In the current study, we sought to determine the neural generators underlying effects of lexical–semantic prediction by using magnetoencephalography (MEG) and event-related potentials (ERPs), supplemented by functional magnetic resonance

imaging (fMRI). In particular, our goal was to localize contextual effects of prediction encouraged by the statistics of the input, as opposed to effects of more passive spreading activation.

Previous work suggests that prediction in language comprehension occurs at multiple levels of representation—syntactic, semantic, lexical, phonological—and often involves the activation of one or more of these representations ahead of the bottom-up input (e.g., Federmeier and Kutas 1999; DeLong et al. 2005; Dikker et al. 2009). Another mechanism that can sometimes result in activation of representations ahead of bottom-up input is automatic spreading activation, which has been proposed to explain the facilitation observed in classic semantic priming paradigms. The relationship between this kind of “passive” priming and sentential prediction and the extent to which they draw on common underlying mechanisms has long been debated (e.g., Neely 1991; Van Petten 1993; Camblin et al. 2007; Chow et al. 2014). What most would agree on, however, is that predictive contexts lead to larger effects on semantic facilitation than passive priming during sentence comprehension (e.g., Van Petten 1993), and that this can offer significant advantages during real-time comprehension: not only does it allow preactivation of stored representations ahead of the bottom-up input, it can also allow these stored representations to be added to the current representation of context such that it is updated in advance of the actual input. For example, prediction may allow the stored memory representation of “cat” to be added to the current message representation after reading the context “She saw a dog chasing a ...” This case can be contrasted with the state of the system after reading the context “She saw a dog ...” Because of the long-term memory connections between “dog” and “cat,” it is likely that “cat” will be passively activated by this context, but it is unlikely that “cat” would be added to the representation of the speaker’s message (see Lau et al. 2013a, for further discussion).

A large electrophysiological literature has demonstrated that both semantic association and contextual predictability are associated with a reduction in a negative deflection in the ERP response to words between 300 and 500 ms, commonly known as the “N400” effect (Kutas and Federmeier 2011). However, when predictability is manipulated by varying the content of the prior context, as in traditional N400 studies of sentence comprehension, it is difficult to determine whether either or both of these mechanisms are responsible for the facilitation observed in “predictive” contexts (Van Berkum, 2009). To avoid this problem in the present study, we examined the effects of prediction on neural activity using a classic semantic priming “relatedness proportion” manipulation, which

allowed us to keep the local context identical across high versus low prediction conditions.

In the relatedness proportion paradigm, predictive validity is varied by manipulating the overall proportion of related prime–target pairs across the course of the experiment (Neely 1976): as the proportion of related prime–target pairs increases, so the size of behavioral priming effects also increases (reviewed by Neely 1991). Since the related prime–target pairs afford the same semantic memory relationships in low-proportion and high-proportion environments, the difference in the size of the priming effect can be attributed to an increase in the extent to which participants use the prime to predict the target. Consistent with this, relatedness proportion usually has little effect on behavioral or neurophysiological priming when the amount of time between prime and target is brief, such that there is insufficient time to generate a prediction (Neely 1976; de Groot 1984; Silva-Pereyra et al. 1999; Grossi 2006; Hutchison 2007; although cf. Bodner and Masson 2003).

It is important to recognize that the relatedness proportion paradigm induces prediction artificially. Therefore, the proximal mechanisms that give rise to the prediction are likely to be quite different from those that support prediction in normal sentence comprehension. However, the strength of this paradigm is that it allows us to probe the “effects” of prediction on lexical–semantic facilitation at 2 extremes—high predictive validity and low predictive validity—while keeping the local context constant across these 2 conditions.

Previous ERP studies using the relatedness proportion paradigm have reported clear evidence of predictive effects over and above passive priming (Holcomb 1988; Brown et al. 2000; Lau et al. 2013a; although cf. Küper and Heil 2010). These studies have largely found that the reduction in the N400 component due to semantic relatedness is significantly larger under conditions of high predictive validity relative to low predictive validity. These results demonstrate that neural activity reflected in the N400 component is modulated by lexical prediction, even when the semantic content of the immediate context is held constant. A similar method has recently been successfully used to demonstrate the existence of a predictive component in neural repetition suppression (Summerfield et al. 2008; Grotheer and Kovács 2014).

Previous work examining the source of N400 effects, using a variety of different paradigms, has mainly implicated 3 areas: left posterior temporal cortex, left anterior temporal cortex, and left inferior frontal cortex (see Van Petten and Luka 2006; Lau et al. 2008, for reviews). fMRI studies using contextual manipulations most frequently show effects in posterior temporal cortex and inferior frontal cortex; some have also reported modulation more anteriorly in temporal cortex (e.g., Rossell et al. 2003), although signal artifact in this region may reduce sensitivity to such differences in fMRI. The limited temporal resolution of fMRI also makes it difficult to unambiguously associate these regions with the ERP effect observed during the N400 time-window. However, it has been argued that effects in the temporal cortex more reliably track the manipulations that elicit N400 effects, and that inferior frontal effects are more likely to reflect postactivation semantic processes that can be modulated by task (Van Petten and Luka 2006; Lau et al. 2008, 2013b). MEG studies using single dipole fits for source localization have usually localized contextual effects in the N400 time-window to midposterior temporal cortex (Helenius et al. 1998; Uusvuori et al. 2008), while MEG studies

using distributed source localization have additionally implicated anterior temporal cortex (Halgren et al. 2002; Lau et al. 2013b) and inferior frontal cortex (Halgren et al. 2002) as generators. Finally, intracranial recordings from anterior inferior medial temporal cortex consistently demonstrate N400-like responses (Nobre and McCarthy 1995; McCarthy et al. 1995), strongly suggesting that this region contributes to the scalp-recorded N400 effect.

The overall pattern of localization results has led several authors to suggest that the N400 component reflects the summed activity of multiple active regions, and that modulations of amplitude can thus arise through multiple generators (Halgren et al. 2002; Pylkkänen and Marantz 2003; Van Petten and Luka 2006). If this is the case, then any differential N400 priming effects observed in high versus low expectation conditions could be due to selective modulation of one of these generators.

To our knowledge, the only previous imaging study to specifically examine the effects of lexical–semantic expectation while holding the context–target relationship constant is a PET study by Mummery et al. (1999), who also used a relatedness proportion semantic priming paradigm. This study revealed effects of relatedness proportion in the left anterior temporal cortex, which, as discussed above, is one of the regions that has been implicated as a source of the N400. However, the use of a block-design analysis in that study made it difficult to distinguish specific effects of predictive facilitation on the target from the processes engaged in actually generating these predictions or attentional reallocation and monitoring mechanisms, which may have been reflected by additional activity observed in anterior cingulate and right superior parietal cortices.

Determining the neural source of reduced activity specifically associated with lexical–semantic prediction is an important first step in better understanding the role of prediction in facilitating neural activity during more natural language comprehension. In the current study, we evaluate which cortical regions specifically contribute to predictive lexical–semantic facilitation, within the N400 time-window (between 350 and 450 ms), by using MEG source localization of the relatedness proportion paradigm, supplemented by data from ERPs and fMRI recorded in the same participants. We recorded electroencephalography (EEG) data simultaneously with MEG in order to better relate contextual effects previously observed in ERPs to the cortical sources implicated by the MEG results. We also recorded fMRI responses in the same paradigm from the same participants with the goal of confirming the results of the MEG source localization.

Materials and Methods

Materials

Each participant took part in 2 separate experimental sessions: one for simultaneous MEG–EEG recordings and one for structural and functional MRI recordings. Therefore, 2 separate sets of stimuli were created. These stimuli have been previously described in Lau et al. (2013a) (The full materials set is available at <http://www.nmr.mgh.harvard.edu/kuperberglab/materials.htm>). In the stimulus list for each session, materials were divided into a low-proportion block in which 10% of word pairs (40/400) were related (low predictive validity), and a high-proportion block in which 50% of word pairs (200/400) were related (high predictive validity). As described below, a core set of controlled test items were counterbalanced across conditions and lists such that each participant saw 40 of these items in each of the 4 conditions, and no participant saw the same word twice. The proportion manipulation was achieved by intermixing these test items with different

proportions of related and unrelated filler pairs and probe filler pairs (see below). In order to prevent item-specific effects, 2 lists were created for each block so that for any given target, half the participants saw the target preceded by a related prime, and half saw the target preceded by an unrelated prime.

To create the related stimuli, 320 highly associated prime–target pairs (e.g., “salt–pepper”) were selected from the University of South Florida Association Norms (Nelson et al. 2004), which were normed by at least 100 participants. All pairs had a forward association strength of 0.5 or higher (at least 50% of participants presented with the prime word responded with the target), with a mean forward association strength of 0.65. The mean log frequency of the targets was 2.55 and the mean log frequency of the primes was 3.53, as computed in the SUBTLEXus (Brysbaert and New 2009). Pairs in which there was clear morphological overlap between prime and target were not included. As the probe task required responding to animal words, no pairs including animal words were included in the test items. These 320 pairs were divided into 2 sets of 160 test items, corresponding to the 2 separate experimental sessions (each containing 40 test items from each of the 4 conditions). Unrelated test items for each set were created by randomly redistributing the primes across the target items, and checking by hand to confirm that this did not accidentally result in any related pairs (e.g., “fuel–pepper”). For each set, test items were then fully rotated across related/unrelated and low-proportion/high-proportion in a Latin Square design across 4 lists, such that no list contained the same prime or target twice. The experimental targets in each list were thus fully counterbalanced across participants (each word could appear in any of the 4 conditions). Forward association strength between prime and target, and log frequency for both prime and target, did not significantly differ between blocks for test items. Each participant saw a completely different set of items from session 1 to session 2, but the order of the sets and the recording modality in which they appeared (EEG–MEG/fMRI) was counterbalanced across participants.

Participants’ task was to respond to an animal word (each animal word appeared exactly once in a session). Therefore, to each block, we added 80 filler unrelated word pairs in which either the prime or target was an animal word (e.g., “jeep–frog” or “hyena–fee”). This was a minor difference from our previous ERP experiment (Lau et al. 2013a) in which animal words always appeared at the target position. This was done in order to ensure that participants paid equal attention to both words of each pair. The primes in the probe trials were never related to the targets.

We achieved the desired relatedness proportion in each block as follows: to the low expectation block, 240 unrelated filler trials (e.g., “lip–museum”) were added to the 80 unrelated probe trials, 40 unrelated trials, and 40 related trials such that, in this block, only 10% of the trials were related. To the high expectation block, 160 related filler trials (e.g., “breakfast–lunch”) and 80 unrelated filler trials were added to the 80 unrelated probe trials, 40 unrelated trials, and 40 related trials, such that, in this block, 50% of the trials were related. The related filler pairs were also selected from the South Florida Association Norms. Because the number of related and unrelated fillers differed across blocks, these items could not be counterbalanced to guard against item-specific effects and are not included in the main analyses here. No word in any position was ever repeated in a given presentation list. The low expectation block was always presented first. Each session was divided into 8 runs of 100 trials. The order of stimuli in each list was randomized using the OptSeq algorithm to improve deconvolution of the hemodynamic response in the parallel fMRI experiment (Dale 1999; <http://surfer.nmr.mgh.harvard.edu/optseq>). The characteristics of the stimuli are summarized in Table 1.

Table 1
Distribution of experimental items across blocks

| | Low proportion | High proportion |
|--|----------------|-----------------|
| Related targets (salt–pepper) | 40 trials | 40 trials |
| Unrelated targets (fuel–pepper) | 40 trials | 40 trials |
| Unrelated animal probes (jeep–frog or hyena–fee) | 80 trials | 80 trials |
| Unrelated fillers (lip–museum) | 240 trials | 80 trials |
| Related fillers (breakfast–lunch) | | 160 trials |

Stimuli were projected onto a screen in white 20-point uppercase Arial font. Each trial began with a fixation cross, presented at the center of the screen for 200 ms, followed by a 200-ms blank screen. The prime word was then presented for 500 ms, followed by a 100-ms blank screen, and then the target word was presented for 900 ms, followed by a 100-ms blank screen. Participants were instructed to press a button on a handheld response box with their left thumb as quickly as possible when they saw a name of an animal. In the MEG session, short intervals were inserted between trials to allow participants time to blink. In the MRI session, additional fixation trials of varying length summing to a total of 60 s per block were added to the intertrial interval in order to optimize deconvolution of event-related activity.

Participants

Thirty young adults initially participated in the MEG–EEG experiment. All participants were native speakers of American English without any prior history of neuropsychiatric disorders, and were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield 1971). Six MEG–EEG datasets were subsequently excluded from the analysis; 5 due to excessive blinking or motion artifact, and 1 due to technical problems with the recording. Analyses were conducted on the remaining 24 datasets (15 men; mean 21 years, range 18–27 years). One of the original 30 participants did not return for the second session of fMRI because of immediately apparent artifact in the initial MEG session. Two of the remaining fMRI datasets were excluded: one due to excessive motion artifact, and one due to a technical error in the stimulus presentation. fMRI analyses were conducted on the remaining 27 datasets (16 men; mean 21 years, range 18–27 years). (After these exclusions, 23 participants had usable data from both recording sessions. A supplementary set of analyses were conducted on this subset of completely overlapping participants. We observed no qualitative differences in these results and the results from the full dataset reported here.) In one of the MEG–EEG sessions and in one of the fMRI sessions, behavioral responses were not recorded due to an equipment error; the behavioral accuracy reported includes the remaining participants. The order of the 2 recording sessions was counterbalanced across participants.

MEG–EEG Recording

The MEG–EEG data were acquired while participants were seated inside a magnetically shielded room (IMEDCO AG, Switzerland). The MEG data were acquired with a dc-SQUID Neuromag VectorView system (Elekta-Neuromag Oy, Finland) with 306 sensors arranged in 102 triplets of 2 orthogonal planar gradiometers and 1 magnetometer. The EEG data were acquired at the same time using a 70-channel MEG-compatible scalp electrode system (BrainProducts, München, Germany) and referenced to an electrode placed on the left mastoid; an electrode was also placed on the right mastoid to confirm that the reference did not incorporate measurable lateralized brain activity. Horizontal and vertical components of eye movements were recorded with 2 pairs of bipolar electrodes placed at the outer canthi and above and below the left eye. Impedance was kept <10 k Ω for all scalp electrode and electrooculogram sites, and <2.5 k Ω for mastoid sites. The EEG and MEG data were acquired with a bandpass of 0.03–200 Hz and were continuously sampled at 600 Hz. To record the head position relative to the MEG sensor array and to co-register the MEG–EEG and MRI coordinate frames, the locations of 3 fiducial points (nasion and 2 auricular), 4 head-position indicator coils, the locations of the EEG electrodes, and at least 100 additional points from the head surface were digitized using a 3 Space Fastrak Polhemus digitizer integrated with the Vectorview system. During the MEG–EEG recording, the head-position coils were used to measure the position and orientation of the head with respect to the MEG sensor array at the beginning of each block of trials.

Structural/Functional MRI Recording

Structural and functional magnetic resonance images were acquired using a 3-T Siemens Trio scanner and a 32-channel head coil. Two T_1 -weighted high-resolution structural images (1-mm isotropic multi-echo MPRAGE: TR = 2.53 s, flip angle = 7°, 4 echoes with TE = 1.64, 3.5, 5.36, 7.22 ms) were acquired at the beginning and end of the session.

To construct the boundary-element model surface for MEG source estimation, a single multiecho 5° flip angle fast low-angle shot (FLASH) image (1 mm isotropic, TR = 20 ms, TE = 1.85 ms + 2*n*, *n* = 0–7), was also acquired.

Eight runs of functional data were collected, each ~5 min long. In each run, 130 functional volumes [36 axial slices (anterior commissure–posterior commissure aligned), 3 mm slice thickness, 0.3 mm skip, 200 mm field of view, in-plane resolution of 3.125 mm] were acquired with a gradient-echo sequence (repetition time = 2 s, echo time = 25 ms, flip angle = 90°, interleaved acquisition).

Preprocessing

Signal-space projection (SSP) correction was applied to MEG magnetometer data to suppress environmental noise and biological artifacts (Uusitalo and Ilmoniemi 1997). For 3 participants in which a strong cardiac artifact was observable in the MEG recording, an additional SSP was computed for gradiometers and magnetometers during heart-beat events (detected using the bipolar ECG electrodes) and applied to the MEG data. Averaged event-related signals time-locked to target words were computed offline from trials free of ocular and muscular artifacts after application of a 20-Hz offline low-pass filter. A 100-ms prestimulus baseline was subtracted from all waveforms prior to statistical analysis. Noisy MEG sensors (0.9% of MEG sensors in the dataset) were marked and excluded prior to generation of grand-average sensor waveforms and source estimates. Prior to EEG sensor analyses, data from noisy or disconnected electrodes (3.4% of electrodes in the dataset) were interpolated with signals from neighboring electrodes using the Laplacian surface estimate.

ERP Analysis

All MEG–EEG analyses reported here focused on the 350–450 ms time-window in which N400 effects of context are at their peak (Kutas and Federmeier 2011). In order to confirm that the relatedness by expectation interaction that we observed in our previous study (Lau et al. 2013a) was replicated in this dataset, we computed a repeated-measures analysis of variance on mean ERP amplitudes between 350 and 450 ms poststimulus onset. In order to assess gross differences in topographical distribution, the analysis was conducted on a subset of 48 nonmidline electrodes divided equally among the 4 quadrants of the scalp, resulting in a 2 × 2 × 2 × 2 design (relatedness × expectation × hemisphere × anteriority). This analysis motivated our planned comparisons in MEG and fMRI, in which we examined the effects of relatedness at each level of expectation.

Individual Source Estimates

The minimum norms estimates (MNE) software package (Gramfort et al. 2014) was used to derive source estimates on the cortical surface for MEG event-related signals in each participant. The cortical surface for each participant was reconstructed from the T₁-weighted MPRAGE structural MRI data using the FreeSurfer software package (<http://surfer.nmr.mgh.harvard.edu>). This high-resolution surface was then decimated into ~10 000 vertices in each hemisphere. A three-compartment boundary-element model with the linear collocation approach was used in the forward calculation (Hämäläinen and Sarvas 1989; Mosher et al. 1999). The scalp and skull surfaces were estimated on the basis of a 5° FLASH structural MRI sequence for most participants; for one participant in which a FLASH sequence was not collected, the MPRAGE structural data were used instead. The amplitudes of the dipoles at each cortical location were estimated for each time sample using the anatomically constrained linear estimation approach (Dale et al. 2000). Noise covariance estimates were derived from data recorded in the 100 ms baseline period prior to the presentation of the prime word for all trials. The orientations of the dipoles were approximately constrained to the cortical normal direction by reducing the variance of the source components tangential to the cortical surface by a factor of 0.5 (Lin et al. 2006a).

Intersubject Registration

Prior to translating individual participant data into a common space, a smoothing operation was applied to the individual data, using 7

iterative steps to spread estimated activity to neighboring vertices. Then each participant's cortical surface was morphed to a template brain created by averaging the cortical surface of 40 individuals scanned by the Buckner laboratory (fsaverage), using an algorithm designed to align individual sulcal–gyral patterns while minimizing distortion (Fischl et al. 1999). The source localization estimates for each participant were mapped, using this same translation, to the template brain space. Quantification of results was then conducted in this template space, and the data from individual participants were averaged in this space for visualization of group-level estimates.

Group-level Analysis of MEG Source Estimates

As described above, dSPM source estimates of activity at the cortical surface for the –100:700 ms surrounding target word presentation were computed for each participant for the 4 conditions of interest (low expectation related, low expectation unrelated, high expectation related, and high expectation unrelated). Next, 2 dSPM contrast maps were created for each participant by subtracting the unrelated–related activity estimates across all vertices, resulting in a dSPM map representing the priming effect for each level of expectation. Then, the activity estimates across the 350–450 ms time-window were averaged in each of these maps to provide an estimate of the mean differential activity in the N400 time-window associated with the priming effect (unrelated–related) for each level of expectation.

Because the aim of this study was to determine which cortical regions specifically contribute to predictive lexical and semantic facilitation, the cortical surface analysis focused on vertices on the cortical surface within areas previously implicated in language processing: bilateral temporal, inferior frontal, inferior parietal, and occipitotemporal cortices (5095 vertices in total), using the Desikan–Killiany atlas (Desikan et al. 2006) included in the FreeSurfer distribution to delineate the regions. We used a nonparametric permutation test based on spatial clustering to estimate which differences in the remaining vertices were reliable across participants (Pantazis et al. 2005; Maris and Oostenveld 2007). For each level of expectation, we first created a *t*-map representing the *t*-statistic associated with the mean priming effect (unrelated vs. related) across the 350–450 ms time-window at each vertex. We established an initial threshold of $t_{(1,23)} = 2.07$ ($P < 0.05$), and grouped the vertices that survived this initial threshold into spatially contiguous clusters. The *t*-values in each cluster were summed, resulting in a second-level cluster statistic. Then, in order to determine whether any of these clusters were not due to chance, we randomly permuted the sign of the priming effect in each subject and ran the same massive univariate *t*-test 1000 times in order to derive a cluster size confidence interval such that $\alpha = 0.05$. Only significant clusters are reported here.

fMRI Analysis

We used FreeSurfer Functional Analysis Stream (FS-FAST) software, with a finite impulse response (FIR) model to analyze the functional MRI data. The FIR model gave estimates of the hemodynamic response on every TR (every 2 s), and thus allowed us to address our hypotheses without assumptions about the shape of the hemodynamic response (Burock et al. 1998; Dale 1999; Burock and Dale 2000). This also allowed us to make a direct comparison with the results of our previous multimodal imaging study of automatic masked semantic priming, which used the same FIR analysis in fMRI. After slice-time correction, functional images were motion corrected to the middle time point of each functional run using the AFNI (afni.nimh.nih.gov/afni) 3 dvolreg program (Cox and Jesmanowicz 1999). Nonbrain voxels were masked out of the analysis using the FSL (www.fmrib.ox.ac.uk/fsl) Brain Extraction Tool (Smith 2002). Images were corrected for temporal drift, normalized, and smoothed using a 3D spatial filter (full-width-half-max: 10 mm). As reviewed above, the cortical surface of each individual was morphed to an average spherical surface representation to align sulci and gyri across subjects (Fischl et al. 1999). This transform was used to map general linear model (GLM) parameter estimates and residual error variances of each participant's functional data to a common spherical coordinate system prior to smoothing. Functional images were then analyzed with a weighted least-square GLM

using the FIR model. Activity estimates between 2 and 12 s poststimulus onset were computed for each condition and summed at each voxel across this time-epoch (Ashby 2011). Estimates were computed only for the 2 cortical surfaces, as MEG is relatively insensitive to subcortical sources and the previous literature has largely implicated cortical areas in semantic processing.

Planned comparisons (unrelated vs. related) were computed for the low expectation and high expectation blocks separately, and assessed using a threshold of $P < 0.05$, FDR corrected. We followed up these comparisons with a more conservative nonparametric test designed to mirror the MEG source analysis procedure. This analysis focused on the same 5095 fronto-temporo-parietal vertices used in the MEG analysis. Significance maps across these vertices were thresholded at $P < 0.05$, uncorrected, and a Monte-Carlo procedure was used to determine cluster-level significance. For each planned comparison, we then set a cluster-level threshold of $P < 0.01$, in order to achieve an effective α of $P < 0.05$ (as both hemispheres and both positive and negative contrasts were tested separately for each planned comparison).

Results

Overall accuracy in detecting the probes (animal primes and targets) in the MEG–EEG session was 90% (SD 0.1%). A 2×2 (expectation \times prime/target) ANOVA revealed a significant main effect of position ($F_{1,22} = 21.6$, $P < 0.01$), reflecting that animal words were detected more accurately in the target position (94%) than in the prime position (85%). There was no main effect or interaction with expectation (F 's < 0.5 , P 's > 0.5), demonstrating that accuracy did not reliably differ between low and high expectation runs.

Consistent with previous work, visual inspection of the ERP (Fig. 1A) suggests the presence of a reduction in N400 amplitude for related relative to unrelated target conditions in both low and high expectation runs, but the effect appears notably greater in the high expectation block than in the low expectation block. This impact of expectation on relatedness was manifested by a significant 3-way interaction between relatedness, expectation, and anteriority ($F_{1,23} = 5.67$, $P < 0.05$), and a significant 4-way interaction between relatedness, expectation, hemisphere, and anteriority ($F_{1,23} = 4.35$, $P < 0.05$). There was also a simple main effect of relatedness ($F_{1,23} = 4.7$, $P < 0.05$) and 2-way interactions between relatedness and hemisphere ($F_{1,23} = 6.54$, $P < 0.05$) and relatedness and anteriority ($F_{1,23} = 5.03$, $P < 0.05$).

In order to further examine the source of the interaction between relatedness and expectation, ANOVAs were conducted within each level of expectation. In the low expectation condition, the numerical trend towards a relatedness effect did not reach significance, and only a marginal 2-way interaction between relatedness and hemisphere was observed ($F_{1,23} = 2.64$, $P = 0.12$; all other P 's > 0.15). In the high expectation condition, however, there was a significant main effect of relatedness ($F_{1,23} = 4.80$, $P < 0.05$), significant 2-way interactions between relatedness and hemisphere ($F_{1,23} = 5.13$, $P < 0.05$) and relatedness and anteriority ($F_{1,23} = 9.10$, $P < 0.05$) and a significant 3-way interaction between relatedness, hemisphere, and anteriority ($F_{1,23} = 5.61$, $P < 0.05$). The interactions with distributional factors in the high expectation conditions reflect the fact that the difference between unrelated and related targets was largest over posterior sites (left posterior: mean difference = 1.58 μ V, right posterior: mean difference = 1.83 μ V) and was particularly small over left anterior sites (left anterior: mean difference = 0.29 μ V, right anterior: mean difference = 1.02 μ V). The right posterior maximum of the N400 effect is consistent with

previous studies using visual presentation (Lau et al. 2013a, see Van Petten and Luka 2006 for review).

For comparison, Figure 1B illustrates the root mean square of activity measured across left and right MEG gradiometers over frontal, temporal, parietal, and occipital cortex. Differences between unrelated and related targets in the 350–450 ms time-window are most apparent in frontal and temporal gradiometers bilaterally, and this effect appears larger for high expectation conditions in temporal sensors.

Cluster-Based Analysis of MEG Distributed Source

Estimates: 350–450 ms

We conducted nonparametric spatial cluster permutation tests on the differential dSPM source estimates for unrelated versus related targets between 350 and 450 ms in the low expectation and high expectation blocks. In the low expectation condition, this procedure revealed no regions that showed a reliable effect of relatedness. For the high expectation condition, this procedure revealed a single region that showed a reliable effect of relatedness. This cluster was in left anterior temporal cortex (Fig. 2), and showed less activity to related targets than unrelated targets. Although differential activity was also visible in corresponding regions of right anterior temporal cortex, these effects did not reach significance.

These results suggest that activity in left anterior temporal cortex contributes to the predictive effects of a semantically related context on facilitating target processing as the likelihood increases that a predictable word will occur. A remaining question is whether left anterior temporal activity is modulated by relatedness only when the target is strongly predicted, or whether the effect of relatedness in this region is simply larger when the target is predicted. The waveform plots in Figure 2 depict the mean source activity estimate from this cluster for all 4 conditions across time. The waveforms in the high expectation condition necessarily diverge in this time-window because the cluster was selected on the basis of this contrast. However, the low expectation conditions demonstrate a similar divergence localized to the same time-window. Although no cluster in the anterior temporal region showed a reliable relatedness effect when expectation was low, the numerical pattern is consistent with the possibility that this region contributed to the small relatedness difference observed in the low expectation condition in ERP. For comparison, Figure 3 illustrates the pattern of mean amplitudes across all 4 conditions in the right posterior quadrant in ERP and the left anterior temporal cluster in MEG.

Region-of-Interest Analysis of MEG Distributed Source

Estimates: 350–450 ms

Effects of contextual facilitation and semantic priming have also frequently been reported in left inferior frontal cortex and left posterior temporal cortex (Van Petten and Luka 2006; Lau et al. 2008). Therefore, although the nonparametric cluster analysis described above did not reveal significant clusters in these areas, we followed up with region-of-interest (ROI) analyses in order to assess the possibility that these regions were also sensitive to expectation. We used the Desikan–Killiany atlas (Desikan et al. 2006) included in the FreeSurfer distribution as a basis for defining the regions, and the posterior temporal ROI was defined by selecting the posterior aspect of the superior temporal sulcus (STS) and middle temporal gyrus (MTG) regions from this parcellation. We conducted a simple

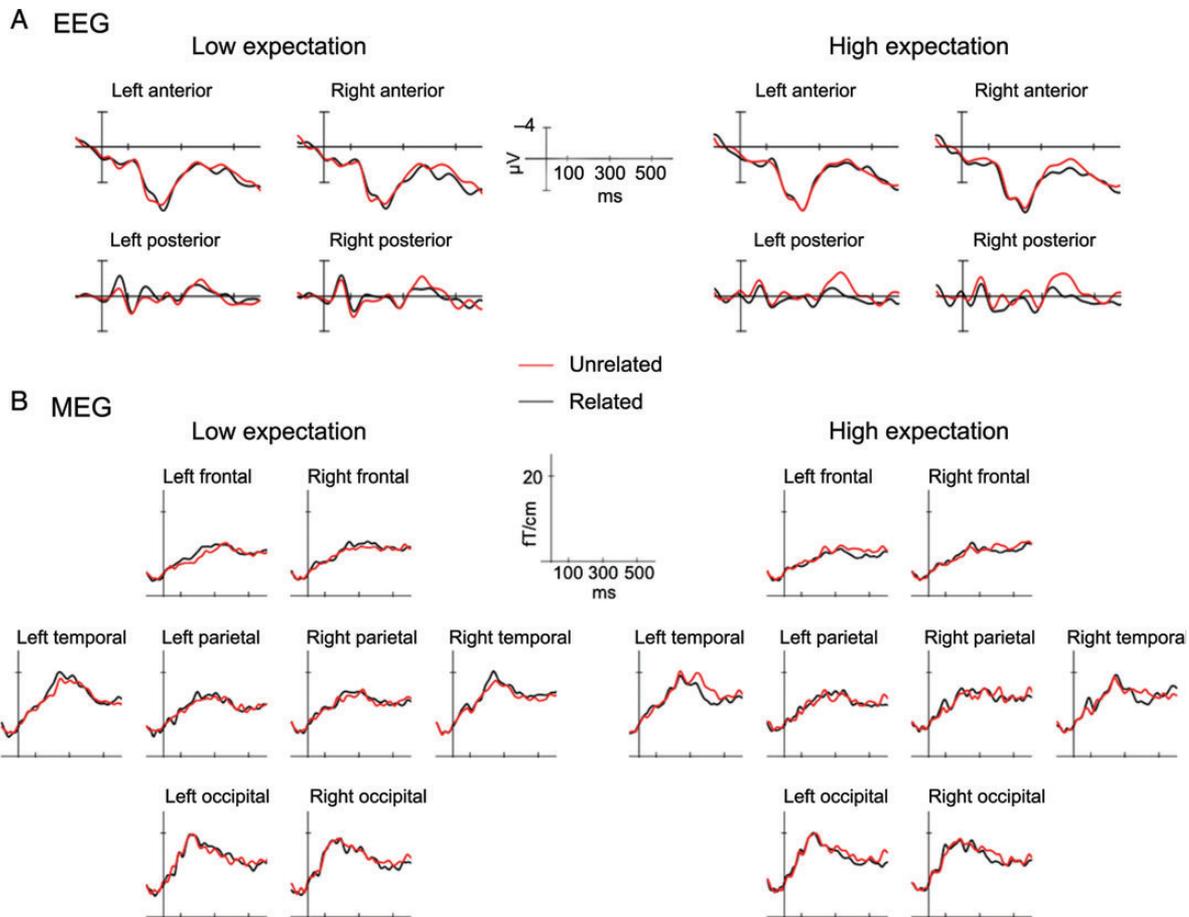


Figure 1. Group-averaged EEG and MEG sensor-level signals, time-locked to the target word, for low and high expectation unrelated minus related contrast. (A) Mean EEG-evoked responses across the 4 quadrant regions in which EEG sensor analyses were conducted. (B) Root mean square MEG-evoked responses, calculated across planar gradiometers only, in frontal, parietal, temporal, and occipital regions.

planned comparison of the relatedness effect (unrelated vs. related) for low and high expectation conditions in each of the 2 ROIs in the 350–450 ms time-window. We found that neither the left posterior temporal ROI nor the left inferior frontal ROI showed significant effects of relatedness in either condition, although there was a numerical trend in the inferior frontal ROI toward a relatedness effect (unrelated > related) in the high expectation condition ($F_{1,23} = 3.0, P = 0.099$; all other P 's > 0.1).

fMRI Analysis

In fMRI, planned comparison of the related and unrelated conditions in the high expectation block was consistent with the MEG source localization results described above by revealing a cluster within the left anterior temporal cortex at an FDR-corrected threshold of $P < 0.05$ (Fig. 2, inset), although this cluster was small and did not survive the more stringent nonparametric Monte Carlo analysis. No such effect was observed in the low expectation conditions at this threshold [The focus of the fMRI FIR analysis in the current paper, which does not assume a canonical hemodynamic response, was primarily to confirm MEG source localization of the N400 predictive facilitation effect (unrelated > related). In other work, we have conducted more extensive fMRI analysis on this dataset, which will be reported in a separate paper (Weber et al. in preparation)].

Discussion

Numerous ERP studies have shown that the amplitude of the N400 is reduced for targets preceded by related primes relative to unrelated primes. More recently, we reported ERP work showing that this N400 reduction is larger when a related prime is more strongly “predicted” (Lau et al. 2013a). In the current study, we not only replicate this finding in ERP, but show with concurrent MEG and parallel fMRI recordings that this predictive modulation is at least partially realized through activity in the left anterior temporal cortex. MEG dSPM source estimates demonstrated a difference between the response to related and unrelated targets in this region during the N400 time-window, that was most prominent (and only reached significance) in the high expectation condition, when the experimental context encouraged participants to predict that the prime and target would be related. fMRI results corroborated this finding and localized the focus of the predictive facilitation effect more precisely to left anterior superior temporal gyrus (STG). Although previous neuroimaging studies have examined the influences of semantic context in comprehension, the unique contribution of the current study is to have isolated the predictive component of contextual facilitation with converging evidence across techniques, while keeping the local context identical across conditions of high and low predictive validity.

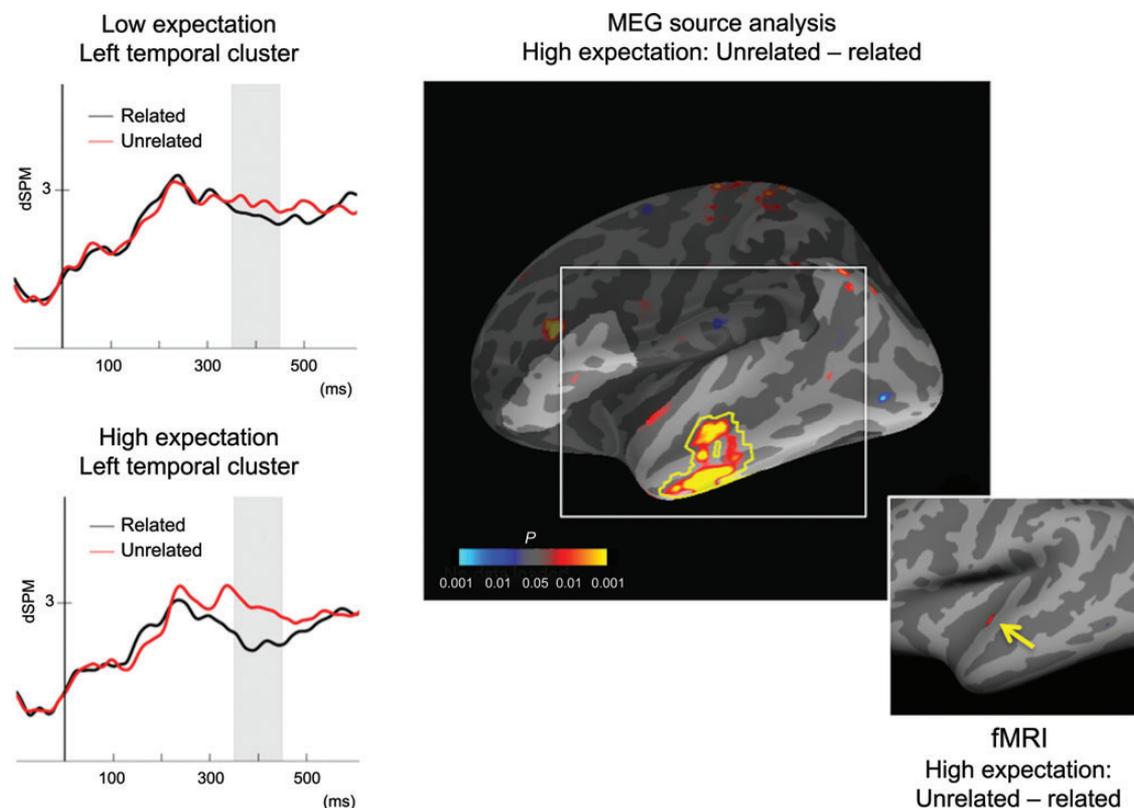


Figure 2. MEG statistical map illustrating t -values for the unrelated-associated contrast of dSPM source estimates between 350 and 450 ms, thresholded at $P < 0.05$. Yellow outline indicates the left anterior temporal cluster that demonstrated a significant reduction for the related condition relative to the unrelated condition in the high expectation block. Lighter shading indicates the search region used in the cluster-level permutation test. Inset: cortical fMRI statistical map illustrating left anterior superior temporal region demonstrating a significant reduction for the related condition relative to the unrelated condition in the high expectation block, FDR corrected at $P < 0.05$. Waveform plots show the mean MEG dSPM activity estimates for the left anterior temporal cluster plotted across the $-100:600$ ms time-window in each condition, for illustration only.

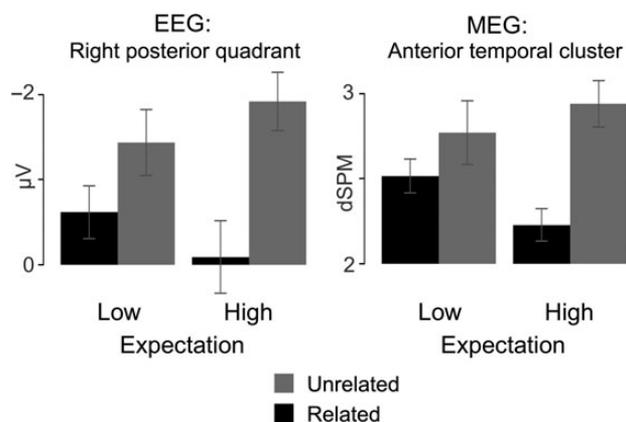


Figure 3. Mean EEG sensor and MEG source amplitudes, respectively, across unrelated and related conditions at each level of expectation for the right posterior EEG sensors at which the N400 effect was at its peak and for the left anterior temporal region demonstrating a significant effect of relatedness in the high expectation block in MEG.

The Role of Temporal Cortex in Predictive Semantic Facilitation

The precise role of left anterior temporal cortex in language comprehension continues to be debated. According to one family of theories, anterior temporal cortex is involved in processing conceptual information through representations that link distributed semantic information from multiple areas

(visual, auditory, motor) to form a “semantic hub” (Mummary et al. 2000; Patterson et al. 2007; Visser et al. 2010; Price 2012). One major source of empirical support from this view comes from cases of semantic dementia, in which damage to anterior temporal cortex bilaterally is correlated with nonlinguistic semantic deficits (Lambon Ralph and Patterson 2008). Our finding that left anterior temporal activity is reduced for words that are strongly predicted is easily accommodated by this view. Many models of lexical-conceptual activation suggest that multiple partially matching representations are initially activated in parallel, and that a “winner” can emerge through competitive interactions in which other representations are inhibited (e.g., McClelland and Elman 1986; Laszlo and Plaut 2012). Prediction of the upcoming word could naturally result in preactivation of the associated conceptual representation, such that when the predicted target is actually presented, the corresponding lexical-conceptual representation more quickly emerges as the winner. If anterior temporal cortex were the locus of the conceptual hub, this network would then be less broadly activated in the highly predicted case, as we observed here.

According to another family of theories, anterior temporal cortex is involved in a secondary, on-the-fly combination of stored lexical or conceptual representations, as suggested by the fact that activity in this region increases when words are presented in phrases or sentences relative to unstructured lists (e.g., Friederici et al. 2000; Vandenberghe et al. 2002; Humphries et al. 2006; Rogalsky and Hickok 2009; Brennan and

Pylkkänen 2012). Consistent with both of these views, much research in speech perception has observed that activity in anterior temporal cortex is correlated with intelligibility (Scott et al. 2000; Spitsyna et al. 2006; Jobard et al. 2007; Lindenberg and Scheef 2007). This theory seems less likely as an explanation for the current findings. This is because, in the present study, with only a few exceptions (e.g., cheddar-CHEESE), most of our associated word pairs could not easily be combined into a well-structured phrase (e.g., salt-PEPPER, shove-PUSH, tardy-LATE, yolk-EGG). Moreover, we saw no modulation of activity in the posterior temporal cortex—an area thought to maintain lexical representations that serve as pointers linking wordform, meaning, and syntactic information (Hickok and Poeppel 2007; Martin 2007)—within the N400 time-window in MEG. On the other hand, it is possible that participants could have engaged in active attempts to combine these items into a phrase (e.g., salt-PEPPER → salt and pepper) or to construct a discourse scenario containing both words of the pair. These kinds of operations would also plausibly be facilitated by lexical-semantic prediction. In other words, during typical sentence comprehension, predicting the next word may not only entail preactivating its stored conceptual representation, but could also facilitate processes of structured combination supported by anterior temporal cortex.

More broadly, it is important to recognize that anterior temporal cortex is a large, functionally heterogeneous region, in which different subparts almost certainly contribute to a variety of semantic processing mechanisms (Bonner and Price 2013). The anterior lateral superior temporal area in which we observed facilitation in both MEG and fMRI is clearly distinct from the medial anterior fusiform region in which intracranial recordings have demonstrated N400-like responses (McCarthy et al. 1995; Nobre and McCarthy 1995). Early PET studies of semantic categorization observed differential activity in anterior inferior temporal gyrus (ITG) (Devlin et al. 2000; 2002). Both MEG and fMRI are subject to signal loss in anterior medial and ventral temporal areas, however, and it may be that we were unable to pick up existing signal in these neighboring areas. The relatively small magnitude of the predictive facilitation effect in fMRI also suggests that the large N400 amplitude differences observed electrophysiologically may have only a moderate impact on the BOLD signal. Finally, in the current study, the exact location of differential anterior temporal activity determined by fMRI and MEG source localization was not identical; in MEG, the cluster showing significant differences was estimated to span anterior STS, MTG, and ITG, while the significant fMRI cluster was located on the insular border of anterior STG. This discrepancy is likely to be due to the limited precision of MEG source localization techniques, as for example inverse methods such as MNE can be biased toward sources that are closer to the sensors (Lin et al. 2006b); but it is also in principle possible that the 2 measures are sensitive to semantic facilitation in different parts of anterior temporal cortex. It will therefore be important for future studies to determine the precise neurocognitive operations subserved by different parts of the anterior temporal lobe, how they interact with one another, and which regions contribute to activity within the N400 time-window.

It will also be important for future studies to dissociate the role of the anterior temporal cortex from a more posterior part of left temporal cortex in which effects of facilitated access have been reported in previous semantic priming paradigms

(Gold et al. 2006; Uusvuori et al. 2008; see Lau et al. 2008 for review), but which was not modulated by prediction in the current study. Several factors may have contributed to the absence of posterior temporal modulation during the N400 time-window. First, in the present paradigm, we used a semantic probe task, which may have elicited greater baseline activity in the conceptual network and thus larger processing benefits due to conceptual prediction. In contrast, lexical decision tasks and word repetition detection tasks such as those used in many of the previous studies may encourage stronger predictions at the lexical level. Second, with respect to previous MEG studies of semantic priming in particular (Uusvuori et al. 2008; Vartiainen et al. 2009), differences in analysis methods may also be responsible for differences in the results observed. In the current study, we estimated distributed solutions of the activity in each condition separately and then focused on the regions that showed significant differences in activity between the 2 conditions. In these previous MEG studies, a single left-hemisphere dipole was selected that best fit activity across all conditions, and the estimated strength of this dipole was shown to differ significantly across conditions. If much of the left temporal cortex is active during lexical processing, but if it is specifically the anterior temporal region that is modulated by semantic prediction, then it could be that the ECD that best explains all activity across conditions falls in the midtemporal cortex, but that the differential activity driving the difference between conditions is mainly due to a slightly more anterior region of temporal cortex.

The Relationship Between Prediction and Passive Priming

As noted in the Introduction, the relatedness proportion manipulation in the current study was designed to artificially emphasize the differential contributions of 2 contextual facilitation processes proposed in the prior literature: passive priming, under conditions of low predictive validity, and prediction, under conditions of high predictive validity. Together with previous work, the current results provide some tentative evidence that the effects of prediction may be realized in the same cortical regions as the effects of passive priming, but to a greater degree.

First, although in the current study, the small N400 effect of relatedness did not reach significance in the low expectation conditions [One factor contributing to the decreased magnitude of the association effect here is likely to be the relatively long stimulus-onset asynchrony (SOA), which is thought to reduce the impact of automatic spreading activation (Neely 1991). However, given that we did observe a low expectation relatedness effect in our previous ERP study using the same materials set (Lau et al. 2013a), we believe the failure to reach significance here can most likely be attributed to the fact that we did not have sufficient power to detect this small effect with the slightly smaller sample size used in the multimodal study ($n = 24$ for EEG-MEG); in the previous study, the effect was just significant ($t = 2.05$) with 32 participants.], the MEG source localization data hint toward the idea that these effects share the same locus as the high expectation effects, as we observed a trend toward an N400 effect in the low prediction conditions in the left anterior temporal cluster in MEG. Second, results from a masked priming paradigm conducted in parallel with the current study (Lau et al. 2013b) indicate that

facilitation due to automatic spreading activation is realized in the same anterior temporal region. In this previous paradigm, ~33% of word pairs were directly related, 33% were indirectly related, and 33% were unrelated. However, as the focus of that study was automatic semantic facilitation, primes were forward and backward masked and presented only 100 ms prior to target onset, resulting in only partial awareness of the prime. Despite the short amount of time available between prime and target, we observed effects of contextual facilitation that were qualitatively similar to the effects observed in the high prediction condition in the current long-SOA paradigm: reductions in left anterior temporal cortex activity observed in fMRI and MEG during the N400 time-window. Furthermore, the location of the peak differential activity in fMRI observed in the current study and the masked priming study (conducted in an overlapping set of participants) are quite similar ($[-50, -14, -4]$ in the current study; $[-52, -9, -3]$ in the masked priming study).

Together, these findings suggest that if automatic spreading activation and prediction are indeed qualitatively distinct mechanisms, they may have the same end result—facilitating conceptual activation in anterior temporal cortex—and point to several important new directions for future work. First, while the current study used word pairs with very strong semantic associates under conditions of fairly high predictive validity (50%) to strongly engage predictive processes, the predictability of natural language contexts is generally less extreme; whereas in the current study, semantically related contexts often predicted a specific word/concept, in natural sentences, context-based predictions are better modeled as a change in the probability distribution over a large set of words or concepts (e.g., [Smith and Levy 2013](#)). Therefore, it will be important for future localization work to determine whether such probabilistic predictions have graded effects on temporal cortex activity. This would be in keeping with theoretical ([Friston 2005](#)) and computational ([Rabovsky and McRae 2014](#)) models in which the effects of passive priming and active prediction are seen as being on a continuum.

Second, while the current study was designed to examine the “effects” of context-based lexical–semantic prediction on semantic facilitation, the process by which predictions are initiated in the relatedness proportion manipulation is likely to be quite different from the initiation of predictions in natural language comprehension. In both cases, comprehenders use the semantic properties of the local context to generate predictions, but in the relatedness proportion paradigm, the global context (the predictive validity of the prime within a block) serves to “gate” these predictions, while natural language involves more complex interactions between semantic associations and higher level syntactic and discourse structure. It will thus be important for future studies to probe the mechanisms by which predictions are generated in more naturalistic contexts; it is worth noting that recent work in sentence and discourse comprehension suggests that global factors such as speaker voice and accent can “gate” predictions of the local linguistic context in a similar fashion ([van Berkum et al. 2008](#); [Hanulíková et al. 2012](#)).

Finally, here we have argued that the contextual effects we observe in the anterior temporal cortex reflect predictive facilitatory effects on semantic activation. However, an alternative account of these results (and all other behavioral and electrophysiological relatedness proportion effects) which does not involve prediction, is that the presence of high relatedness proportion in some way induces a global amplification in the state

of all connections in the lexical–semantic network, such that spreading activation between all lexical–semantic nodes is increased above normal levels during the high relatedness proportion block (or conversely, that spreading activation is reduced below normal levels during the low relatedness proportion block). Although we do not think this can be ruled out as an explanation for relatedness proportion effects, we are not aware of existing semantic network models that have the property that the strength of all connections can be globally altered on the fly. It is also not clear how such a model could explain other properties of relatedness proportion effects demonstrated in previous work, for example, the reduced prevalence of such effects when the lag between prime and target is brief ([Neely 1976](#)). These and other findings have motivated the widely held conception of spreading activation as an automatic process that proceeds in the same way regardless of higher level context.

Conclusion

In sum, in this study we isolated the contribution of predictive processes to the facilitation effect associated with semantic priming by manipulating the proportion of related word pairs across the experiment. The MEG results show that experimental conditions of increased predictive validity lead to reduced activity in left anterior temporal cortex in response to predicted words. Coupled with simultaneously acquired ERP data showing that prediction was associated with an increased N400 reduction, these results support the hypothesis that facilitation due to prediction is an important component of the N400 context effect. They also converge with previous results in suggesting that this effect is driven in part by modulation of activity in anterior temporal cortex. Taken together with previous work implicating the left anterior temporal cortex in semantic processing, our results suggest that activation of the semantic network can be more narrowly constrained as the likelihood increases that the word can be predicted by prior context, resulting in an overall reduction of neural activity in this network 350–450 ms after word onset. These results suggest that predictive paradigms provide a promising research strategy for more precisely pinpointing the functional role of anterior temporal cortex in language comprehension.

Funding

This work was supported by the National Institutes of Health (grants R01MH071635 to G.R.K., F32HD063221 to E.L., and 5R01EB009048 and P41RR14075 to M.H.), as well as a grant from the Ellison Foundation to M.H.

Notes

We thank Scott Burns, Phillip Holcomb, Candida Ustine, Sheraz Khan, Nao Suzuki, Naoro Tanaka, Seppo Ahlfors, Kana Okano, Nandita Shetty, Sam Mehl, Anna Namyst, and Ben Stillerman for valuable assistance. *Conflict of Interest:* None declared.

References

- Asby FG. 2011. Statistical analysis of fMRI data. Cambridge (MA): MIT Press.
- Bodner GE, Masson ME. 2003. Beyond spreading activation: an influence of relatedness proportion on masked semantic priming. *Psychon Bull Rev.* 10(3):645–652.

- Bonner MF, Price AR. 2013. Where is the anterior temporal lobe and what does it do? *J Neurosci*. 33(10):4213–4215.
- Brennan J, Pyllkkänen L. 2012. The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage*. 60(2):1139–1148.
- Brown CM, Hagoort P, Chwilla DJ. 2000. An event-related brain potential analysis of visual word priming effects. *Brain Lang*. 72(2): 158–190.
- Brysbaert M, New B. 2009. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav Res Meth*. 41(4):977–990.
- Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM. 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport*. 9(16): 3735–3739.
- Burock MA, Dale AM. 2000. Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. *Hum Brain Mapp*. 11(4):249–260.
- Camblin CC, Gordon PC, Swaab TY. 2007. The interplay of discourse congruence and lexical association during sentence processing: evidence from ERPs and eye tracking. *J Mem Lang*. 56(1):103–128.
- Chow WY, Lago S, Barrios S, Parker D, Morini G, Lau E. 2014. Additive effects of repetition and predictability during comprehension: evidence from event-related potentials. *PLoS One*. 9(6):e99199.
- Cox RW, Jesmanowicz A. 1999. Real-time 3D image registration for functional MRI. *Magnet Reson Med*. 42:1014–1018.
- Dale AM. 1999. Optimal experimental design for event-related fMRI. *Hum Brain Mapp*. 8:109–114.
- Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E. 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*. 26:55–67.
- de Groot AMB. 1984. Primed lexical decision: combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *Q J Exp Psychol Hum Exp Psychol*. 36(2):253–280.
- DeLong KA, Urbach TP, Kutas M. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat Neurosci*. 8(8):1117–1121.
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31(3):968–980.
- Devlin JT, Russell RP, Davis MH, Price CJ, Moss HE, Fadili MJ, Tyler LK. 2002. Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia*. 40(1):54–75.
- 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(6):589–600.
- Dikker S, Rabagliati H, Pyllkkänen L. 2009. Sensitivity to syntax in visual cortex. *Cognition*. 110(3):293–321.
- Federmeier KD. 2007. Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology*. 44(4):491–505.
- Federmeier KD, Kutas M. 1999. A rose by any other name: long-term memory structure and sentence processing. *J Mem Lang*. 41(4):469–495.
- Fischl B, Sereno MI, Tootell RBH, Dale AM. 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*. 8:272–284.
- Friederici AD, Meyer M, von Cramon DY. 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang*. 74:289–300.
- Friston K. 2005. A theory of cortical responses. *Phil Trans R Soc B*. 360(1456):815–836.
- Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH. 2006. Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *J Neurosci*. 26(24):6523–6532.
- Gramfort A, Luessi M, Larson E, Engemann D, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen MS. 2014. MNE software for processing MEG and EEG data. *Neuroimage*. 86:446–460.
- Grossi G. 2006. Relatedness proportion effects on masked associative priming: an ERP study. *Psychophysiology*. 43(1):21–30.
- Grotheer M, Kovács G. 2014. Repetition probability effects depend on prior experiences. *J Neurosci*. 34(19):6640–6646.
- Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, Dale AM. 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*. 17(3):1101–1116.
- Hämäläinen MS, Sarvas J. 1989. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans Biomed Eng*. BME-36:165–171.
- Hanulíková A, van Alphen PM, van Goch MM, Weber A. 2012. When one person's mistake is another's standard usage: the effect of foreign accent on syntactic processing. *J Cogn Neurosci*. 24(4):878–887.
- Helenius P, Salmelin R, Service E, Connolly JF. 1998. Distinct time courses of word and context comprehension in the left temporal cortex. *Brain*. 121(6):1133–1142.
- Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8(5):393–402.
- Holcomb PJ. 1988. Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain Lang*. 35(1):66–85.
- Humphries C, Binder JR, Medler DA, Liebenthal E. 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J Cogn Neurosci*. 18(4):665–679.
- Hutchison KA. 2007. Attentional control and the relatedness proportion effect in semantic priming. *J Exp Psychol Learn Mem Cogn*. 33(4):645–662.
- 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.
- Küper K, Heil M. 2010. Letter search and relatedness proportion: further electrophysiological evidence for the automaticity of semantic activation. *Neurosci Lett*. 482(1):26–30.
- Kutas M, Federmeier KD. 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu Rev Psychol*. 62(1):621–647.
- Lambon Ralph MA, Patterson K. 2008. Generalization and differentiation in semantic memory. *Ann N Y Acad Sci*. 1124(1):61–76.
- Laszlo S, Plaut DC. 2012. A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain Lang*. 120(3):271–281.
- Lau E, Stroud C, Plesch S, Phillips C. 2006. The role of structural prediction in rapid syntactic analysis. *Brain Lang*. 98(1):74–88.
- Lau EF, Gramfort A, Hämäläinen MS, Kuperberg GR. 2013b. Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *J Neurosci*. 33(43):17174–17181.
- Lau EF, Holcomb PJ, Kuperberg GR. 2013a. Dissociating N400 effects of prediction from association in single-word contexts. *J Cogn Neurosci*. 25(3):484–502.
- Lau EF, Phillips C, Poeppel D. 2008. A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci*. 9(12):920–933.
- Lin FH, Belliveau JW, Dale AM, Hämäläinen MS. 2006a. Distributed current estimates using cortical orientation constraints. *Hum Brain Mapp*. 27:1–13.
- Lin FH, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW, Hämäläinen MS. 2006b. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum norm estimates. *Neuroimage*. 31(1):160–171.
- Lindenberg R, Scheef L. 2007. Supramodal language comprehension: role of the left temporal lobe for listening and reading. *Neuropsychologia*. 45(10):2407–2415.
- Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods*. 164:177–190.
- Martin A. 2007. The representation of object concepts in the brain. *Annu Rev Psychol*. 58:25–45.
- McCarthy G, Nobre AC, Bentin S, Spencer DD. 1995. Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. *J Neurosci*. 15(2):1080–1089.
- McClelland JL, Elman JL. 1986. The TRACE model of speech perception. *Cogn Psychol*. 18(1):1–86.

- Mosher JC, Leahy RM, Lewis PS. 1999. EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng.* 46:245–259.
- Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RSJ, Hodges JR. 2000. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann Neurol.* 47(1):36–45.
- Mummery CJ, Shallice T, Price CJ. 1999. Dual-process model in semantic priming: a functional imaging perspective. *Neuroimage.* 9(5):516–525.
- Neely JH. 1976. Semantic priming and retrieval from lexical memory: evidence for facilitatory and inhibitory processes. *Mem Cogn.* 4:648–654.
- Neely JH. 1991. Semantic priming effects in visual word recognition: a selective review of current findings and theories. In: Besner D, Humphreys GW, editors. *Basic processes in reading.* Hillsdale (NJ): Erlbaum. p. 264–336.
- Nelson DL, McEvoy CL, Schreiber TA. 2004. The University of South Florida free association, rhyme, and word fragment norms. *Behav Res Methods Instrum Comput.* 36(3):402–407.
- Nobre AC, McCarthy G. 1995. Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *J Neurosci.* 15(2):1090–1098.
- Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia.* 9(1):97–113.
- Pantazis D, Nichols TE, Baillet S, Leahy RM. 2005. A comparison of random field theory and permutation methods for the statistical analysis of MEG data. *Neuroimage.* 25:383–394.
- Patterson K, Nestor PJ, Rogers TT. 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci.* 8:976–987.
- Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage.* 62(2):816–847.
- Pylkkänen L, Marantz A. 2003. Tracking the time course of word recognition with MEG. *Trends Cogn Sci.* 7(5):187–189.
- Rabovsky M, McRae K. 2014. Simulating the N400 ERP component as semantic network error: insights from a feature-based connectionist attractor model of word meaning. *Cognition.* 132(1):68–89.
- 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.
- Rossell SL, Price CJ, Nobre AC. 2003. The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia.* 41(5):550–564.
- Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain.* 123(12):2400–2406.
- Silva-Pereyra J, Harmony T, Villanueva G, Fernández T, Rodríguez M, Galán L, Díaz-Comas L, Bernal J, Fernández-Bouzas A, Marosi E et al. 1999. N400 and lexical decisions: automatic or controlled processing? *Clin Neurophysiol.* 110:813–824.
- Smith NJ, Levy R. 2013. The effect of word predictability on reading time is logarithmic. *Cognition.* 128(3):302–319.
- Smith SM. 2002. Fast robust automated brain extraction. *Hum Brain Mapp.* 17(3):143–155.
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJ. 2006. Converging language streams in the human temporal lobe. *J Neurosci.* 26(28):7328–7336.
- Summerfield C, Trittschuh EH, Monti JM, Mesulam M-M, Egner T. 2008. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci.* 11(9):1004–1006.
- Uusitalo MA, Ilmoniemi RJ. 1997. Signal-space projection method for separating MEG or EEG into components. *Med Biol Eng Comput.* 35(2):135–140.
- Uusvuori J, Parviainen T, Inkinen M, Salmelin R. 2008. Spatiotemporal interaction between sound form and meaning during spoken word perception. *Cereb Cortex.* 18(2):456.
- Van Berkum JJA. 2009. The neuropragmatics of ‘simple’ utterance comprehension: an ERP review. In: Sauerland U, Yatsushiro K, editors. *Semantic and pragmatics: from experiment to theory.* Basingstoke, UK: Palgrave MacMillan. p. 276–316.
- Van Berkum JJA, Van Den Brink D, Tesink CM, Kos M, Hagoort P. 2008. The neural integration of speaker and message. *J Cogn Neurosci.* 20(4):580–591.
- Vandenberghe R, Nobre AC, Price CJ. 2002. The response of left temporal cortex to sentences. *J Cogn Neurosci.* 14(4):550–560.
- Van Petten C. 1993. A comparison of lexical and sentence-level context effects in event-related potentials. *Lang Cogn Proc.* 8(4):485–531.
- Van Petten C, Luka BJ. 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain Lang.* 97(3):279–293.
- Van Petten C, Luka BJ. 2012. Prediction during language comprehension: benefits, costs, and ERP components. *Int J Psychophysiol.* 83(2):176–190.
- 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 MA. 2010. Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci.* 22:1083–1094.
- Wicha NY, Moreno EM, Kutas M. 2004. Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *J Cogn Neurosci.* 16(7):1272–1288.