

Brain Networks Subserving the Extraction of Sentence Information and Its Encoding to Memory

Uri Hasson^{1,2}, Howard C. Nusbaum^{2,3,4} and Steven L. Small^{1,2,3,4}

¹Department of Neurology, ²Department of Psychology, ³The Center for Cognitive and Social Neuroscience and ⁴The Brain Research Imaging Center, The University of Chicago, Chicago, IL, USA

Sentences are the primary means by which people communicate information. The information conveyed by a sentence depends on how that sentence relates to what is already known. We conducted an fMRI study to determine how the brain establishes and retains this information. We embedded sentences in contexts that rendered them more or less informative and assessed which functional networks were associated with comprehension of these sentences and with memory for their content. We identified two such networks: A frontotemporal network, previously associated with working memory and language processing, showed greater activity when sentences were informative. Independently, greater activity in this network predicted subsequent memory for sentence content. In a separate network, previously associated with resting-state processes and generation of internal thoughts, greater neural activity predicted subsequent memory for informative sentences but also predicted subsequent forgetting for less-informative sentences. These results indicate that in the brain, establishing the information conveyed by a sentence, that is, its contextually based meaning, involves two dissociable networks, both of which are related to processing of sentence meaning and its encoding to memory.

Keywords: deactivation, default network, encoding, individual differences, language, semantic memory

Introduction

Language is used to communicate information, and sentences are the primary means by which information is communicated. However, the information conveyed by a sentence depends on the relation between the sentence and what is already known. For example, the sentence “the neighbor’s dog bit his owner” is less informative (LI) if the dog had a history of biting people than if the dog was known to have a gentle manner. Language comprehension entails understanding the information conveyed by sentences in a given context, and retaining that information.

Our goal in this research was to identify the brain networks involved in 2 functions: Establishing the information sentences convey in relation to context and encoding that information. In an fMRI study, we presented volunteers with spoken sentences (e.g., “the dog bit his owner”) that conveyed either more or less information depending on the preceding context. We identified 2 sorts of brain regions: 1) those implicated in establishing the information conveyed by sentences *in context* and 2) those involved in encoding the content of these sentences to memory. Our main objective was to identify brain regions associated with these 2 functions so that we could establish whether, on the neurophysiological level, the 2 functions are subserved by separate brain networks or by a single network.

Brain regions involved in the online extraction of sentence information were expected to show greater activity during processing of sentences when those sentences were more informative (MI) because these express messages that are less probable given the prior context (e.g., Shannon 1948; Bar-Hillel 1964).¹ Prior research has identified a number of regions where neural activity varies with sentences’ contextually induced meaning: For example, narratives have been found to evoke greater activity than unlinked sentences in regions including the anterior temporal lobe (aTL) bilaterally, temporoparietal junction (bilaterally), the left middle frontal operculum (BA 45), left middle temporal gyrus (MTG), left superior temporal sulcus (STS), medial prefrontal cortex, premotor and motor cortex (e.g., Giraud et al. 2000; Vogeley et al. 2001; Xu et al. 2005). In particular, 2 studies (Ferstl et al. 2005; Virtue et al. 2006) suggest that the dorsomedial prefrontal cortex (dmPFC), the left inferior frontal gyrus (IFG), and lateral temporal regions are specifically involved in elaboration of sentences’ content in relation to context. In the study by Ferstl et al., participants listened to narratives and judged whether their endings were consistent or inconsistent with prior context. Inconsistent endings evoked greater activity in the anterior IFG (more extensively on the left), right aTL, and the dmPFC. In the study by Virtue et al., participants heard actions that were described explicitly or implicitly (e.g., John needed to prepare for a wedding; he went to “change/find” clothes). Implicit actions, which rely on access to conceptual knowledge, evoked greater activity in the right superior temporal gyrus (STG). Later references to these actions (e.g., John emerged with a tuxedo) evoked greater activity in the left STG and MTG when referring to implicit actions (left IFG activity was found only for individuals with higher working memory). The relevance of these studies extends beyond identifying neural systems mediating comprehension of sentences in context. Specifically, the difference between their results suggests that such systems are sensitive to the task participants are engaged in: Virtue et al. who used a passive task report left temporal activity and no dmPFC activity. In contrast, Ferstl et al., who used an explicit analytic task that demanded monitoring texts for consistency, report extensive frontal activity, no left temporal activity, and more anterior right temporal activity. Engagement in strategic tasks such as this has been shown to affect activity in regions associated with speech comprehension (e.g., Binder et al. 2004; Blumstein et al. 2005). Furthermore, in regions associated with language comprehension, such tasks can increase differences between experimental conditions that vary in processing difficulty (e.g., left IFG; Hasson et al. 2006; Rodd et al. 2005). Given that our main interest was in investigating those processes that accompany spontaneous language comprehension, we

employed a passive listening task. Therefore, we expected that sentences more informative in context would be associated with increased activity in temporal regions, and possibly also in dmPFC and left IFG.

Brain regions involved in the encoding of sentences to memory were expected to show greater neural activity during processing of sentences that are subsequently remembered (vs. sentences that are subsequently forgotten). Certain brain regions such as the left IFG have been implicated in encoding stimuli into memory because neural activity in those regions during stimuli processing predicts whether or not these stimuli will be subsequently remembered (Paller and Wagner 2002). We expected that these regions would also be associated with memory for sentence content. However, we were particularly interested to know if the networks involved in the encoding of sentences into memory are sensitive to the informativeness of those sentences. Specifically, sentences that introduce more information may be more difficult to encode, and so the increased neural activity typically associated with successful subsequent memory would vary as a function of sentences' informativeness.

By independently identifying regions sensitive to the informativeness of sentences during online comprehension and regions involved in the encoding of sentence content, we could determine whether these form a single functional network. Some researchers have argued that encoding into memory is a specialized function (e.g., Tulving 2001), whereas others have contended that semantic processing and memory encoding are 2 facets of the same process because certain brain regions are involved in both (e.g., Fletcher et al. 2003; Otten et al. 2001, 2002; Otten and Rugg 2001a; Wagner et al. 1998). For example, neuroimaging studies have found that when individuals perform judgments about single words (e.g., is "tiger" animate or inanimate), left IFG shows increased activity during judgments that are more semantically complex, and concurrently, neural activity in the same region predicts which words will be subsequently remembered (cf., Paller and Wagner 2002; Wagner et al. 1998). Such findings have led to the suggestion that left IFG is part of a semantic working memory network where semantic elaboration results in more effective encoding of materials, at least at the single-word level (Gabrieli et al. 1998). On the basis of prior studies, we therefore expected that IFG should differentiate between more- and less-informative sentences and that in this region neural activity during sentence comprehension would be predictive of subsequent memory for sentences.

However, it is important to note that our study departed from the aforementioned studies of subsequent memory in that those employed single printed words, whereas ours employed elaborate sentences presented in a spoken discourse context. This difference was likely to produce different patterns of neural processing because sentences presented in discourse contexts evoke greater activity in lateral temporal regions than do single words (e.g., Xu et al. 2005). Similarly, spoken stimuli evoke more temporal activity than printed stimuli (e.g., Michael et al. 2001; Constable et al. 2004). Indeed, sentences have been repeatedly shown to strongly engage lateral temporal regions (e.g., Humphries et al. 2001; Friederici 2002; Vandenberghe et al. 2002) and we therefore expected to find strong activity in these regions.

Whether temporal activity would be associated with subsequent memory was less clear. On the one hand, prior studies examining the neural correlates of subsequent memory for

single printed words have not consistently reported an association between activity in temporal regions and subsequent memory: Some have found an association (e.g., Otten et al. 2002; Reber et al. 2002; Uncapher and Rugg 2005) and some have not (e.g., Otten and Rugg 2001a; Fletcher et al. 2003). On the other hand, some research suggests that activity in temporal regions could predict subsequent memory for complex narratives. First, Casasanto et al. (2002) asked participants to memorize short context-independent sentences and examined the neural correlates of interindividual differences in memory for those sentences. Participants with better memory showed greater neural activity in the posterior MTG, supramarginal gyrus, and IFG (all left hemisphere) during sentence presentation. Second, behavioral and imaging studies of causal processing suggest that temporal regions could be involved in encoding. Behavioral studies have shown that texts expressing cause-effect relations of intermediate strength are better remembered than texts where such relations are either very strong or nonexistent (e.g., Keenan et al. 1984). Complementing these behavioral findings, 2 imaging studies (Mason and Just 2004; Kuperberg et al. 2006) report that in certain brain regions, texts expressing intermediate-strength causal relations evoke greater neural activity than texts expressing very strong or nonexistent relations. Whereas the 2 studies report different loci for this effect (Mason and Just: right temporal regions, Kuperberg et al.: no right temporal regions, but left middle temporal, bilateral inferior frontal, bilateral angular gyrus, bilateral medial and superior frontal gyrus, among others), the correspondence between the imaging data and the behavioral memory findings suggests that these regions are involved in establishing causal links between the sentences, which results in better memory for those sentences. Thus, temporal, inferior frontal, and superior frontal regions could be implicated in both the extraction of sentence information and its subsequent memory.

In our investigation of these issues, participants passively listened to short stories during an fMRI scan. These 8-sentence stories were constructed so that target sentences near the end of each story (the sixth sentence) differed in their informativeness as a function of preceding context (see Table 1). That is, prior context determined whether the events described in the target sentences were more or less likely to occur. A surprise forced-choice recognition test was given after the scan that enabled us to identify which stories were remembered and forgotten for each participant. Thus, we could evaluate which regions were sensitive to informativeness in context and which were associated with subsequent memory for discourse content.

Methods

Participants

Twenty-three participants (10 men and 13 women, mean age 21.4 years, SD = 2.6) were recruited from the student population of The University of Chicago. All were right handed and had normal hearing and normal (corrected) vision. The study was approved by the Institutional Review Board of the Biological Science Division of The University of Chicago, and all participants provided written informed consent.

Stimulus Construction

The experimental materials consisted of 40 pairs of short (8-sentence) spoken narratives. In each pair, the 2 stories were identical, apart from 1 or 2 words in the third sentence that changed the context in such a way that altered the informativeness of a later (sixth) sentence that was the

Table 1

Examples of meaningful and filler experimental materials

Meaningful story, example 1

Tom has ordered new dinner plates from Mikasa
 As always, he placed the order using the Internet
 The plates are scheduled to arrive within 3 to 5 days [MI]/weeks [LI]
 Tom's previous dinner plates were of poor quality
 They chipped quickly and Tom wanted better plates
 His Internet order arrived after a month
 Tom was happy to receive the order
 He cannot remember the last time he entered a department store

Meaningful story, example 2

At IBM, chips are coated in a section called "the clean room"
 Technicians use special chemicals to coat the chips
 The FDA labels these chemicals as safe [MI]/toxic [LI]
 IBM closely tracks its workers' safety
 Its most recent report was published yesterday
 It reported that working in the clean room caused chemical burns
 IBM is studying plans to deal with this issue
 Its workers' safety is its major concern

Example of filler sequence

She puts the ink to her right
 She packs her entire wardrobe in one trunk
 He sits down with his friends at a booth
 Bill is a stunt expert who works in Hollywood
 Sam decides to pawn off his valuables
 He spends every morning and afternoon in the library
 People put glasses on the tray for refills
 She pushes the keyboard off the desk

target of our analysis (see Table 1 for examples). Therefore, although these target sentences were identical in the 2 experimental conditions, in one condition they were less informative and in the other they were more informative, with relation to the comprehender's representation of the preceding discourse. Behavioral (e.g., Albrecht and O'Brien 1993; O'Brien et al. 1998), ERP (e.g., van Berkum et al. 1999, 2003) and imaging research (Ferstl et al. 2005) have shown that similar manipulations affect the ease of integrating a sentence with prior context. We normed the materials to ascertain that the MI endings were less expected given the prior context: We presented the pairs of stories to a group of volunteers ($n = 10$) who did not participate in the study, and asked them to indicate, for each pair, in which of the 2 story versions was the ending less expected. In 96% of the trials, participants chose those versions we referred to as MI. For 21 of the 40 story pairs, all participants gave responses that confirmed to our categorization (i.e., stating that the MI condition was less expected), and for each of the remaining 19 stories, no less than 8 (i.e., 8 or 9) participants agreed with our categorization. The 40 story pairs were assigned to 2 experimental lists so that each participant heard either the LI or the MI version of each particular pair. In addition to the 40 experimental items, each list included 20 filler items. The filler materials included meaningful sentences that did not amount to a coherent narrative (see Table 1). The order of condition types, that is, filler, LI, and MI trials in each list, was established using software that determined the presentation order of the experimental conditions so that the resulting design was optimized for estimation of the unknown parameters when using regression-based deconvolution (RSFgen routine; <http://afni.nimh.nih.gov/afni>). Once the order of the 60 trial types (20 LI, 20 MI, 20 filler) was determined, we created 2 complementary experimental lists of trials in which the stories appearing in their LI form at any given position in one list appeared in the identical position in their MI form in the other list. The assignment of specific story to position in the lists was done randomly. After removing 3 subjects whose data could not be used in the analysis, 10 subjects were assigned to one list and 10 to the other. The excluded subjects made too few errors (none or one) in the recognition test (for the LI or MI conditions) and thus their data could not be included in within-participant statistical analyses comparing accurate to inaccurate recognition. Each experimental list was partitioned into 8 runs presented consecutively: 7 runs of 8 items each and 1 run of 4 items.

Sentences were recorded to digital tape by a male speaker in a sound attenuated recording booth, and converted to computer files (16-bit stereo, 44 kHz sampling rate). The sentences' volume was mean

normalized, and the sentences concatenated to stories. Each sentence was between 2.5 and 3.5 s long, and the interval between the onset of 2 consecutive sentences was exactly 4 s. Each story (or filler item) was preceded by a 2-s orientation tone followed by 8 sentences (32 s) and a break (20 s). We included these relatively long breaks after each of the 60 auditory texts so that we could reliably assess how activity during language comprehension related to activity in the absence of an exogenous task, as explicated below.

Procedure

Inside the scanner, participants first completed a short volume calibration stage: They were presented with sentences while the scanner emitted the sounds associated with a functional scan. Participants iteratively indicated via gesture whether they wanted to increase or decrease the volume level until a comfortable level was achieved. The instructions stated in part: "In this study, you will be listening to stories, which will be delivered over headphones. Your task is to follow the stories presented over the headphones and attend to their contents. At the end of the study you will be asked some general questions about your impressions of the stories. Each story is quite short, around half a minute long, and each is followed by a short break ... Every once in a while, a series of sentences will be presented instead of a story. In such cases you are to simply understand those sentences." Participants listened to the stories passively (i.e., without secondary task requirements), and approximately 15 min after the fMRI scan, their memory for all 40 experimental stories was assessed with a forced-choice recognition task. Memory for filler materials was not assessed.

Behavioral Data Collection

In the recognition test that followed the scan, we told participants that we were interested in seeing which stories left an impression on them, and that we assessed this by having them read pairs of stories and decide which of the 2, if any, they had heard during the experiment. We emphasized that this was not a test of intelligence, and that they could take as long as they needed to make their decision. Participants were also told that depending on the version of the experiment they were assigned to, the computer could display one or more pairs of stories in which neither of the stories had been presented in the scanner; in reality, no such pairs were presented. These instructions were intended to reduce the possibility that participants would arbitrarily choose one of the 2 stories even when they did not remember either of the 2. One of the stories was the one participants heard in the scanner and the other, the lure story, was the matching story from the other condition. These stories appeared one above the other (counterbalanced across trials), and participants pressed "1," "2," or "3" to indicate they had recognized hearing the upper story, lower story, or none of the two.

Image Acquisition and Preprocessing

Scans were acquired on a 3-Tesla scanner using spiral acquisition with a standard head coil. Two volumetric T_1 -weighted scans (120 axial slices, $1.5 \times 0.938 \times 0.938$ mm resolution) were acquired and averaged to provide high-resolution images on which to identify anatomical landmarks and onto which functional activation maps could be superimposed. For the functional scans, 32 spiral T_2^* gradient echo images covering the entire brain were collected every 2 s in the axial plane (time repetition = 2 s; time echo = 30; flip angle = 77). Effective functional resolution was 3.8 mm^3 . We collected 1620 whole-brain images (216 in each of the first 7 runs and 108 in the last run). Images were spatially registered in 3-dimensional space by Fourier transformation of each of the time points and corrected for head movement, using AFNI. For each participant, the raw signal in each voxel was scaled to the mean of the voxel's signal during the study. Time points associated with extreme head movement were removed from the regression models (<1% of data).

fMRI Data Analysis

Story contents were conditionalized separately for each participant on the basis of their subsequent memory performance. This resulted in 5 sorts of experimental conditions: MI-correct, MI-mistake, LI-correct, LI-mistake, and filler.

Estimations of signal intensity in each condition on the individual-subject level were followed by second-level group analyses. Intensity values in the functional imaging data were analyzed using multiple linear regression. Regressors were waveforms with similarity to the hemodynamic response, generated by convolving a gamma-variant function with the onset time and duration of the blocks of interest. One regressor was modeled to reflect the initial 5 sentences in the experimental conditions, and the other regressors reflected whether the final 3 sentences belonged to the MI-correct, MI-mistake, LI-correct, LI-mistake, or filler conditions. Additional regressors were the mean, linear, and quadratic trend components, as well as the 6 motion parameters in each of the functional runs.

For the second-level group analyses, we aligned the participants' functional data to a common space by inflating each hemisphere of the anatomical volumes to a surface representation and aligning it to a template of average curvature using the FreeSurfer software package (Dale et al. 1999; Fischl et al. 1999). The resulting representations of surface curvature were imported into SUMA (Saad et al. 2004) that is a software that enables surface mapping of functional data (<http://afni.nimh.nih.gov/afni/suma/>). We used SUMA to project the functional data (i.e., regression parameter estimates) from the 3-dimensional volumes onto the 2-dimensional surfaces. This procedure results in accurate reflection of the individual data at the group level (Argall et al. 2005). Following, participants' data were smoothed on the surface tessellation with a Gaussian 4-mm FWHM filter to decrease spatial noise. Smoothing the data on the surface rather than on the volume avoids inclusion of white matter data in the result of the smoothing, and also avoids averaging of data across sulci (e.g., Desai et al. 2005; Kuperberg et al. 2006). Analyses were conducted using the AFNI software package and the "R" statistical software package (<http://www.r-project.org/>).

We conducted the group level statistical analyses after projecting data from the volume domain to the surface domain. All group level statistical analyses were thresholded to control for a family wise error rate (FWE) of $P < 0.05$ unless noted otherwise. Threshold parameters were determined by Monte Carlo simulations (Forman et al. 1995) using AFNI's ALPHASIM routine. These simulations control for FWE by estimating what volume an activation cluster needs to exceed to be considered as reliable. The relevant parameters for these simulations are the inter-voxel correlation (which increases cluster size) and the intensity that should hold for each voxel in the cluster (specified as a lower bound P value).

To determine which regions showed significant task-related activation in both the LI and the MI conditions, for each condition we identified regions where activity (signal change) was positively correlated with the task (individual voxel threshold, $P < 0.001$, $P < 0.05$ corrected), and considered only those voxels that were part of reliable clusters in both conditions; regions where activity was negatively correlated with the task (deactivation) in both conditions were similarly established (reported in Table 3). Note that *activation* and *deactivation* are defined by whether a voxel's time series was positively or negatively correlated with regressors that consisted of modeled idealized hemodynamic response functions. To graphically depict active regions (white outline in Figs 2 and 3), we identified regions showing significant activation separately in each condition (individual voxel threshold, $P < 0.05$, $P < 0.05$ corrected), and considered only those voxels that were part of reliable clusters in both cases. A more expository analysis contrasting the filler condition with the MI and LI conditions is reported in Supplementary Figure S1 (individual voxel threshold, $P < 0.01$, uncorrected).

To determine which regions showed a context effect, we contrasted the MI and LI contents remembered correctly (i.e., tested for MI-correct \neq LI-correct; an analysis including all items revealed the same patterns, but somewhat less reliably and is available upon request from the authors). We considered that the experimental effects might be manifested in small and isolated clusters where all voxels survive a strict threshold (localized activations), or larger clusters where all voxels survive a lower threshold (diffuse activations). To this end, we conducted 2 analyses that equally controlled for FWE, $P < 0.05$. In the one probing for localized clusters, we set the individual voxel threshold at $P < 0.001$, and simulations indicated that a reliable cluster would exceed 9 contiguous voxels (486 mm^3). In the one probing for regional clusters, we set the

individual voxel threshold at $P < 0.05$ and simulations indicated that a reliable cluster would exceed 73 contiguous voxels (3492 mm^3).

To examine subsequent memory effects, we entered the intensity values into a 2 (memory: recalled, forgotten) \times 2 (context: LI, MI) analysis of variance and identified regions showing a main effect of subsequent memory (recalled \neq forgotten) or an interaction between the 2 factors. We examined both localized and diffuse clusters. Reliable clusters exceeded 11 contiguous voxels when the single voxel threshold was set at $P < 0.001$, and exceeded 86 contiguous voxels when the single voxel threshold was set at $P < 0.05$ (both tests, FWE $P < 0.05$, corrected).

The analysis of the correlation between participants' overall memory performance and their neural activity in the LI and MI conditions was conducted by parcellating each brain into anatomical regions of interest (ROIs). All surface ROIs were delineated using automatic parcellation methods (Fischl et al. 2004) in which the statistical knowledge base derives from a training set incorporating the anatomical conventions of Duvernoy (1991). The accuracy of these methods has been shown to be similar to that of manual parcellation (Fischl et al. 2002, 2004). Our hypotheses were focused on 6 regions in the temporal and frontal lobes, with each hemisphere considered separately (for a total of 12 regions): 1) STG (lateral aspect, not including the supratemporal plane or Heschl's gyrus), 2) STS, 3) MTG; and the 3 subdivisions of the IFG: 4) pars opercularis (\sim BA 44), 5) pars triangularis (\sim BA 45), and 6) pars orbitalis (\sim BA 47). We chose these regions because, as discussed in Introduction, they are the regions most commonly implicated in both single-sentence and discourse-level processes. (Other regions sometimes implicated in these processes—medial and inferior parietal regions—were also analyzed and are reported in Supplementary Table S1.) For each region, we established the correlation between participants' mean blood oxygen level-dependent signal across all voxels in that region and their memory performance. The motivation for this brain-behavior correlation analysis was derived from the results of the behavioral data, which indicated strong interindividual variability in attention to the materials.

To identify regions that showed both a context effect and a subsequent memory effect we conducted 2 analyses. One analysis identified voxels that belonged to regional clusters in both the analysis of context effects and the analysis of subsequent memory effects. To this end, we overlaid the independently thresholded statistical map of regions showing a subsequent memory effect (remembered $>$ forgotten, $P < 0.05$ corrected) onto the independently thresholded statistical map of regions showing an effect of informativeness in context (MI-correct $>$ LI-correct, $P < 0.05$ corrected; cf., Nichols et al. 2005). To ensure that the overlap was not a result of one of the 4 experimental conditions driving both main effects, we removed from the resulting map all regions showing an interaction between the 2 factors. The second analysis probed for more localized regions associated with both effects by identifying voxels that were independently reliable at a threshold of $P < 0.01$ in each of the 2 analyses (joint probability, $P < 0.0001$), with a cluster extent of at least 10 voxels (534 mm^3). Again, we removed from the analysis any voxel showing a reliable interaction between the 2 factors.

To determine which voxels showed a correlation between the magnitude of the context effect and that of the subsequent memory effect, we conducted a voxel-wise analysis in those regions where both effects were reliable. In this analysis, for each voxel we correlated the magnitude of the 2 effects across participants to establish Pearson's product moment correlation coefficient (r) and assessed the correlation's significance using a t -test where $t = r \times \sqrt{(N-2)/(1-r^2)}$, and t has a Student t distribution with $N-2$ df (i.e., 21 dfs). The individual voxel threshold was set at $P < 0.005$ ($r = 0.59$). Simulations indicated that to control for multiple comparisons, clusters would need to exceed 12 voxels. A similar analysis was conducted to identify clusters where all voxels showed a correlation between neural activity in the LI or MI conditions and participants' recognition accuracy. This analysis was a whole-brain analysis with simulations conducted after projection of the data to the surface domain. The single voxel threshold was set at $P < 0.005$ ($r = 0.59$), and simulations (following procedures in Nichols and Holmes 2002) indicated that reliable clusters would need to exceed 473 surface vertexes ($\sim 0.25\%$ of total number of vertexes in a hemisphere's surface area).

Results

Behavioral Results

Following the imaging experiment, participants completed a forced-choice recognition test in which we assessed which experimental stories were remembered or forgotten on an individual-participant basis (see Methods). Participants' mean accuracy was 70% (SD = 17); individual accuracy ranged from 33% to 98% correct (14/40 and 39/40 stories, respectively). The distribution of responses across the story types and the relationship between participants' overall accuracy and the types of error they made are given in Table 2 and Figure 1.

Table 2 shows that response accuracy was similar for stories with MI and LI endings ($P > 0.25$ in an analysis by subjects, $P > 0.44$ in an analysis by items) and that the distribution of errors was quite similar across the 2 conditions. An odds-ratio analysis revealed no reliable difference between the distribution of errors across the 2 conditions (odds ratio = 1.13; the 95% confidence interval included 1 and ranged between 0.69 and 1.9). However, the distribution of these 2 sorts of errors across participants was associated with participants' overall accuracy (Fig. 1). Participants who made more errors made a larger proportion of errors as a result of indicating that neither story was heard (Pearson's $r = 0.45$, $P < 0.05$). This correlation indicated that participants with low and high accuracy may have paid different degrees of attention to the materials which prompted us to conduct an analysis of interindividual differences reported later.

fMRI Results

Our analytic approach to data analysis consisted of 4 steps. First, to verify our basic results against the prior literature, we id-

Table 2

Performance on the forced-choice memory recognition test following the study (+SE)

Condition	Correct identification	Error: choosing lure	Error: indicating neither story was presented
MI	72% (2)	13% (1)	15% (2)
LI	68% (2)	16% (1)	16% (2)

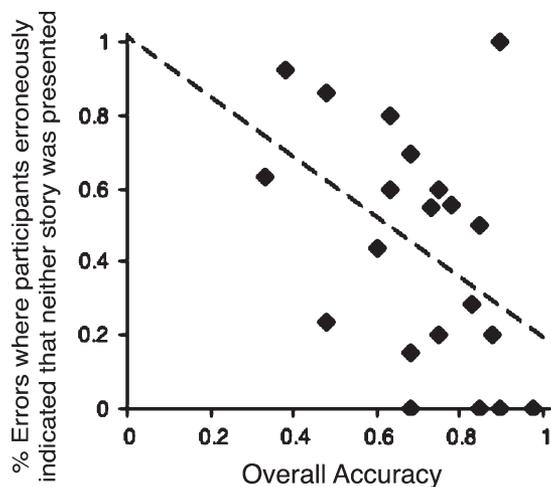


Figure 1. Individual performance on forced-choice memory test. The scatter diagram plots each participant's overall performance in the forced-choice memory test against the proportion of errors where they wrongly indicated that neither story was presented.

entified brain regions that showed either above- or below-baseline neural activity during language comprehension. These were defined in reference to the 20 sec rest intervals that followed each story (activation vs. deactivation henceforth). The following steps addressed the main theoretical questions at the basis of the current study: In the second step, we analyzed regions that showed differential processing to story-ending segments as a function of their informativeness in context (i.e., MI vs. LI in context; henceforth an effect of context). In the third step, we assessed in which regions neural activity predicted subsequent memory for discourse contents (a subsequent memory effect) or showed a differential subsequent memory effect as a function of context (a context \times subsequent memory interaction). In the fourth and final step, we joined the results of the former 2 analyses to identify those regions that independently demonstrated both sensitivity to informativeness in context and a subsequent memory effect.

Activation and Deactivation during Language Comprehension

In this analysis, we probed for regions showing activation or deactivation in both the MI and the LI conditions, so that we could compare the basic results against prior literature. The results are summarized in Table 3. Reliable activation for story ending segments in both conditions was found in regions typically implicated in spoken language comprehension (e.g., bilateral temporal regions, the left inferior and superior frontal gyri; e.g., Skipper et al. 2005; Hasson et al. 2006). Reliable deactivation in both experimental conditions was found in portions of dorsal prefrontal cortex (ventral and medial), the inferior parietal lobule, and large clusters in midline regions (precuneus and posterior and anterior cingulate). These patterns of deactivation are remarkably similar to those previously reported in the literature for a large variety of cognitive tasks including auditory ones (e.g., Shulman et al. 1997; McKiernan et al. 2003). Thus, our basic findings for neural activity during spoken language comprehension proved in accordance with prior literature. A secondary analysis revealed that the MI and LI conditions showed greater activity than the filler condition in temporal regions (bilaterally), indicating that these regions

Table 3

Regions showing reliable activation (or deactivation) for both more- and less-informative story-ending segments

Region	Volume	x	y	z	BA
Brain regions showing reliable activation in both more informative and less informative conditions					
L. STG	38154	-47	-13	-1	22
R. STG	26932	53	-21	4	41
L. SFG	3420	-8	35	53	6
L. MFG/IFG	2298	-46	24	24	46
R. MTG	641	47	-57	22	39
L. Cingulate	588	-8	13	39	32
L. MFG	428	-34	47	-5	10
Brain regions showing reliable deactivation in both more informative and less informative conditions					
L. Precuneus	21001	-4	-64	38	7
R. Cingulate	10367	6	24	36	32
R. IPL/SMG	8336	44	-41	43	40
R. Cingulate	4649	1	-22	34	24
L. MFG	4328	-28	40	34	9
L. IPL/SMG	3046	-52	-36	38	40
R. MFG	2512	34	29	40	8
L. MFG	1282	-25	-1	58	6

Note: Individual voxel threshold $P < 0.001$, FWE $P < 0.05$, for each condition (i.e., joint probability of $P < 0.00001$). Center of mass identified by Talairach and Tournoux coordinates.

were sensitive to the narrative structure of the materials (Supplementary Figure S1).

Neural Correlates of Sentence Informativeness in Context

We examined which regions showed differential activity to the informativeness of story-ending segments as a function of preceding context (i.e., MI vs. LI conditions). We considered that the effect of prior context could be manifested in regional (diffuse) clusters of activity or localized clusters of activity and probed for both types (both controlled for FWE at $P < 0.05$ corrected for whole-brain comparison, see Methods). The results of both analyses are presented in Figure 2 (centers of mass for localized clusters are given in Table 4). As the figure shows, neural activity during processing of a story-ending segment was greater when the segment was more informative in context than when the exact same segment was less informative in context.

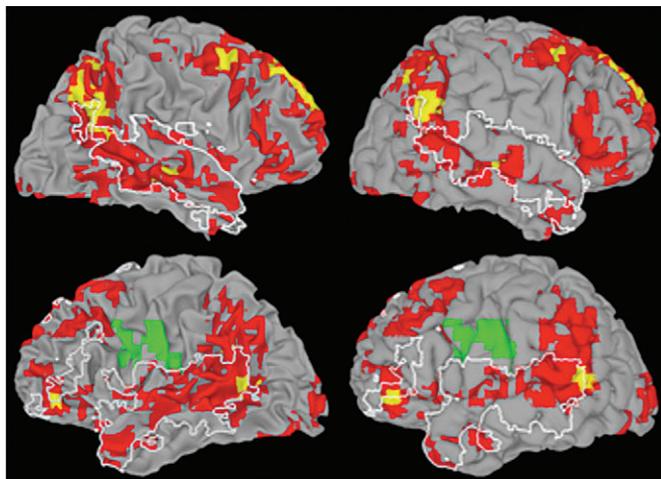


Figure 2. Effects of context. Data are projected onto white matter and gray matter views of standard Montreal Neurological Institute template. Comprehension of concluding discourse segments was generally associated with greater neural activity when those segments were more informative (yellow: voxel threshold $P < 0.001$, $P < 0.05$ corrected; red: voxel threshold $P < 0.05$, $P < 0.05$ corrected; green: regions showing greater activity in the LI condition, voxel threshold $P < 0.05$, $P < 0.05$ corrected). Regions showing above-baseline activity in both conditions are marked in white outline (see text).

Table 4

Regions showing greater blood oxygen level-dependent signal when story-ending segments were more informative than when the same segments were less informative

Region	Volume	x	y	z	BA
R. SFG	8871	11	46	38	8
L. Cingulate (posterior)	7962	-2	-52	28	31
R. STG	4756	45	-55	26	39
L. MTG	1550	-51	-53	8	39
R. Thalamus	1389	1	-12	11	
Brainstem	1122	1	-29	-5	
L. Cingulate (central)	1015	0	-16	40	24
R. MFG	962	37	6	49	6
R. Lingual G.	908	2	-75	-1	18
L. IFG	748	-50	36	1	45
L. Supramarginal G.	695	-46	-45	24	40
R. STG	641	57	-23	-1	21
R. SFG	641	10	23	60	6
R. Superior parietal lobule	534	37	-56	48	7/40

Note: Individual voxel threshold $P < 0.001$, corrected $P < 0.05$. Center of mass identified by Talairach and Tournoux coordinates.

In temporal and inferior parietal areas, increased activity in the MI condition was found bilaterally, from inferior parietal areas most posteriorly, extending anteriorly along STG, STS, and MTG to the temporal poles. The right IFG showed increased activation in all 3 subdivisions of the gyrus (pars opercularis, triangularis, and orbitalis), whereas left IFG showed increased activity mainly in the anterior part of IFG (pars orbitalis). There was also bilateral activity in the middle and superior frontal gyri (MFG, SFG), dmPFC, and posterior midline regions (not shown in figure). Interestingly, regions in the left precentral and postcentral gyri demonstrated greater activation for the LI condition.

There was a substantial overlap between regions that showed general activity during spoken language comprehension (i.e., above-baseline activity for both MI and LI conditions; delimited by white outline in Fig. 2) and those that showed greater activity in the MI condition: Of the total volume of regions that demonstrated reliable activation for both conditions, 44% overlapped with reliable clusters showing more activity in the MI condition. The pattern of overlap was particularly revealing for the left IFG that showed increased sensitivity to sentences' informativeness along the posterior-anterior axis (see Fig. 2): The pars opercularis and part of the pars triangularis showed reliable activity for both MI and LI conditions, but did not differentiate between them. More anteriorly, a posterior aspect of pars orbitalis (BA 47) demonstrated above-baseline activity in both conditions but also increased activity in the MI condition. Finally, the most anterior aspect of pars orbitalis demonstrated greater activity in the MI condition, without demonstrating above-baseline activity for both conditions. Reliably greater activity in the MI condition was also found in dmPFC bilaterally, but more so on the right (Fig. 2). We discuss the findings for IFG and dmPFC in Discussion.

Some of the clusters showing context effects (14% of total clusters' volume) overlapped with regions showing deactivation in both MI and LI conditions; these were mostly found in inferior parietal regions (bilaterally), the precuneus, and to a lesser extent in right frontomedial regions. We discuss the findings in the deactivation network after reviewing the results of the memory analysis.

To summarize, we found that regions often implicated in sentence-level processing were sensitive to the informativeness of sentences in a given context, with increased neural activity found when sentences were more informative in context. These effects were not limited solely to temporal and inferior frontal regions usually associated with language comprehension, but extended to inferior parietal, prefrontal, and midline regions as well.

Neural Correlates of Memory: Subsequent Memory Effects

In this analysis, we examined in which regions neural activity predicted subsequent memory for story contents. Following previous studies (e.g., Schott et al. 2006), for each participant we partitioned the stories as a function of whether their contents were subsequently remembered or forgotten and whether the story ending was more or less informative. This procedure resulted in 4 story types: MI-correct, MI-mistake, LI-correct, and LI-mistake. As in the previous analysis, we probed for regional and localized clusters of activity. Regions showing subsequent memory effects are given in Figure 3 and center of mass coordinates for localized clusters are provided in Table 5.

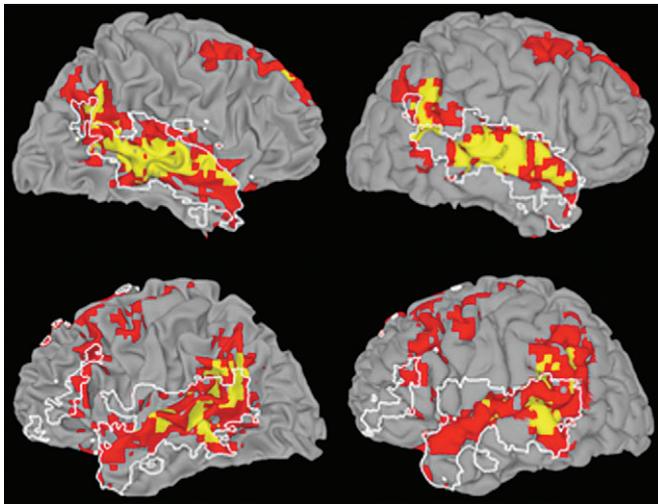


Figure 3. Correlates of subsequent memory for discourse. Data are projected onto white matter and gray matter views of standard Montreal Neurological Institute template. Successful subsequent memory was associated with greater neural activity that was found in both regional clusters (red: voxel threshold $P < 0.05$, $P < 0.05$ corrected) and more localized clusters (yellow: voxel threshold $P < 0.001$, $P < 0.05$ corrected). Regions showing above-baseline activity in both conditions are marked in white outline (see text).

Table 5

Regions where blood oxygen level-dependent signal during comprehension of story endings was associated with subsequent memory

Region	Volume	x	y	z	BA
Subsequent memory effect ^a					
R. STG	16031	53	-31	8	22
L. STG	8764	-53	-40	12	22
L. STG	1122	-52	-16	5	22/41
R. SFG	962	21	39	37	9
R. STG	748	54	7	-6	38
L. Insula	695	-39	-6	-7	13
Interaction between subsequent memory and context ^b					
R. Medial FG	2672	11	44	34	9
L. Precuneus	1924	-7	-56	46	7
R. Precuneus	1603	2	-63	33	6/7
R. STG	641	48	13	-11	38

Note: Individual voxel threshold $P < 0.001$, corrected $P < 0.05$. Center of mass identified by Talairach and Tournoux coordinates.

^aIn all cases, these regions demonstrated increased activity for those materials later remembered correctly.

^bIn all cases the interaction reflected $(MI\text{-correct} - MI\text{-mistake}) > (LI\text{-correct} - LI\text{-mistake})$.

Replicating previous studies (e.g., Wagner et al. 1998; Otten et al. 2001, 2002; Otten and Rugg 2001a; Fletcher et al. 2003; Uncapher and Rugg 2005), we found that accurate subsequent memory was associated with increased neural activity during task performance. Regional clusters showing a subsequent memory effect included the left IFG (pars triangularis), bilateral inferior parietal and temporal regions extending from the supramarginal gyrus and the posterior-dorsal STG/MTG to the temporal poles, and prefrontal regions. Localized clusters were found in the right dmPFC and bilaterally in temporal regions.

Crucially, in some regions the magnitude of the subsequent memory effect depended on the informativeness of the story-ending segment (a context \times subsequent memory interaction), indicating that these regions were involved in memory encoding and were also sensitive to the informativeness of sentences.

In all cases, the interaction reflected a larger subsequent memory effect for the MI condition (i.e., $[MI\text{-correct} - MI\text{-mistake}] > [LI\text{-correct} - LI\text{-mistake}]$). Localized clusters were found in the right dmPFC, right STG, and the precuneus bilaterally (Fig. 4 and Table 5). To understand the interaction patterns, we considered each cluster as a functional ROI, and for each we obtained the mean neural activity in each of the 4 conditions (vs. the resting period; see Fig. 4). A similar analysis was conducted for regional clusters: These overlapped to a certain extent with the localized clusters but were also found in left midline regions, and bilaterally in inferior parietal and frontal regions (Fig. 5). As Figures 4 and 5 show, the patterns of neural activity were remarkably similar across regions. In general, less deactivation was associated with better memory in the MI condition but worse memory in the LI condition. This novel and extremely consistent pattern suggests that patterns of neural activity that are associated with a return to a “default mode” of activity (i.e., resting baseline) can be associated either with facilitated or with impeded subsequent memory. We elaborate on this issue in the Discussion.

Neural Correlates of Memory: Individual Differences

The error patterns in the postscan recognition test suggested that strategic differences might be responsible for different levels of accuracy among participants (see Fig. 1). Such differences presented us with the opportunity to understand which brain regions mediate interindividual differences in memory performance. Consequently, we conducted a between-participants correlation analysis to identify regions where neural activity during comprehension of the MI and LI story endings correlated with participants' overall recognition accuracy. Note that performance in the recognition test reflected whether participants successfully integrated the meaning of the target sentence with prior context during the comprehension of the stimuli. Previous experimental work has shown that this particular ability depends on general comprehension skills (Long and Chong 2001). Given that our prior analyses showed involvement of lateral temporal regions in both hemispheres in establishing sentence information in relation to context, we wanted to establish whether relatively poor memory performance in some participants was associated with reduced activity in these regions. Further, because IFG has been associated with memory formation in prior studies (cf., Paller and Wagner 2002) but did not strongly show this relation in our prior analysis, we wanted to specifically probe the relation between activity in this region and interindividual differences in subsequent memory.

It is important to note the difference between the current interindividual analysis and the one reported in the section Neural Correlates of Memory: Subsequent Memory Effects. Whereas that analysis identified regions that differentiated remembered materials from forgotten materials for the participants as a group (technically, participants were modeled as a random factor), the current between-participants correlation analysis examines whether neural activity in certain anatomical regions systematically differs between participants as a function of their performance. These regions may or may not coincide with those demonstrating subsequent memory effects in the previous analysis.

For this analysis, we used automatic parcellation tools (Fischl et al. 2004) to delineate inferior frontal and temporal regions that have often been implicated in language comprehension (as reviewed in Introduction), and in each we examined the correlation between the participants' memory accuracy and their

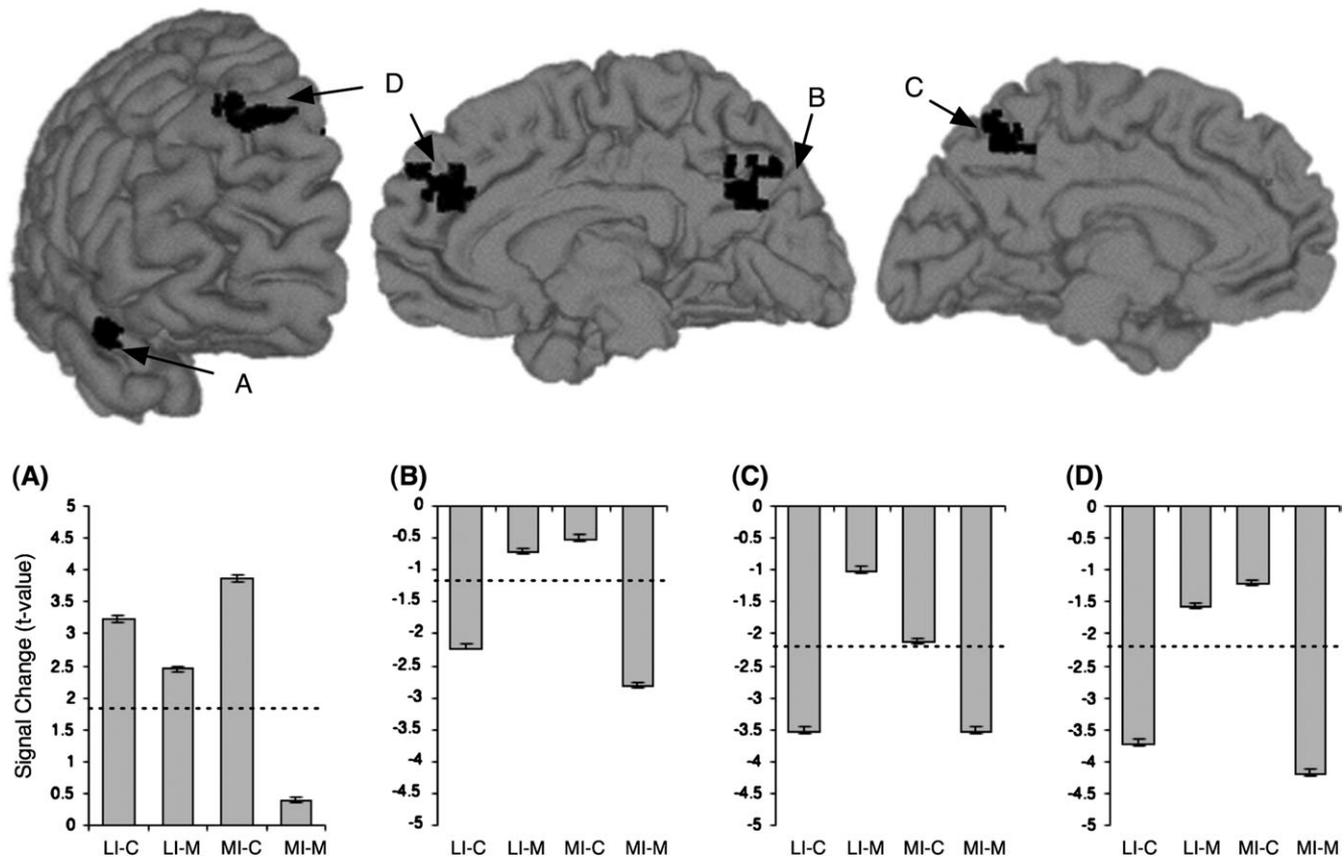


Figure 4. Localized clusters showing an interaction between discourse context and memory performance. These clusters (voxel threshold $P < 0.001$, $P < 0.05$ corrected) correspond to (A) the R. STG, (B) R. precuneus, (C) L. precuneus, and (D) R. medial frontal gyrus. LI-C = less-informative correct, LI-M = less-informative mistake, MI-C = more-informative correct, MI-M = more-informative mistake. Error bars show standard error of the mean calculated over voxels. Dashed lines mark the mean activity in each cluster for the filler condition, which consisted of meaningful sentences that did not make up a coherent narrative.

neural activity during comprehension of story endings in the MI and LI conditions. These regions (bilaterally) consisted of the 3 subdivisions of the IFG (pars opercularis, triangularis, orbitalis), the STG, STS, and MTG. As shown in Figure 6, this analysis revealed reliable associations between neural activity and individual memory accuracy in the left pars orbitalis (in the MI condition) and bilaterally in the temporal regions. No reliable associations were found in pars triangularis or opercularis. An exploratory analysis that also included medial, inferior parietal, and dorsal frontal regions revealed only one additional region where activity was predictive of memory accuracy; this was found for the right superior frontal gyrus in the MI condition (Supplementary Table S1).

This analysis extends the previous subsequent memory analysis in that it demonstrates that temporal regions not only show differential activity for forgotten and remembered items but also show a strong association between overall neural activity during comprehension and participants' performance in the forced-choice recognition test. In left IFG, this pattern held only for the left pars orbitalis that also demonstrated greater activity in the MI condition than in the LI condition. This analysis, however, lacks the fine-grained resolution required to identify which aspects of the temporal cortex were the loci of such correlations. We therefore conducted a complementary whole-brain voxel-wise analysis that identified clusters where all voxels showed a reliable correlation between participant's neural activity and recognition accuracy (see Methods). This analysis

identified such correlations for both the LI and MI conditions in the middle third of the right STS, primarily on the inferior bank, and one cluster in the anterior third of the left STS that showed such correlations in the MI condition (Supplementary Figure S2). No other brain regions were identified in this analysis.

Regions Showing Both an Effect of Informativeness in Context and a Subsequent Memory Effect

To assess the presence of simultaneous context and subsequent memory effects, we overlaid the independently thresholded statistical map of regions showing a subsequent memory effect onto the independently thresholded statistical map of regions showing an effect of informativeness in context (see Methods). As seen in Figure 7A, a joint effect of informativeness and subsequent memory was evident in temporal and middle frontal regions. Although the left IFG was implicated in both effects, their respective loci did not extensively overlap. Overlaps were also found in inferior parietal regions (on the left) and SFG (bilaterally). To summarize the results we also identified more localized clusters of overlapping functionality (Table 6). In this analysis, we included voxels that independently survived an individual voxel threshold of $P < 0.01$ in both the MI > LI contrast and the remembered > forgotten contrast and set a cluster threshold of at least 10 voxels. This analysis revealed clusters of activity in STG bilaterally, right prefrontal cortex, and the left insula. Areas of overlap largely excluded regions showing

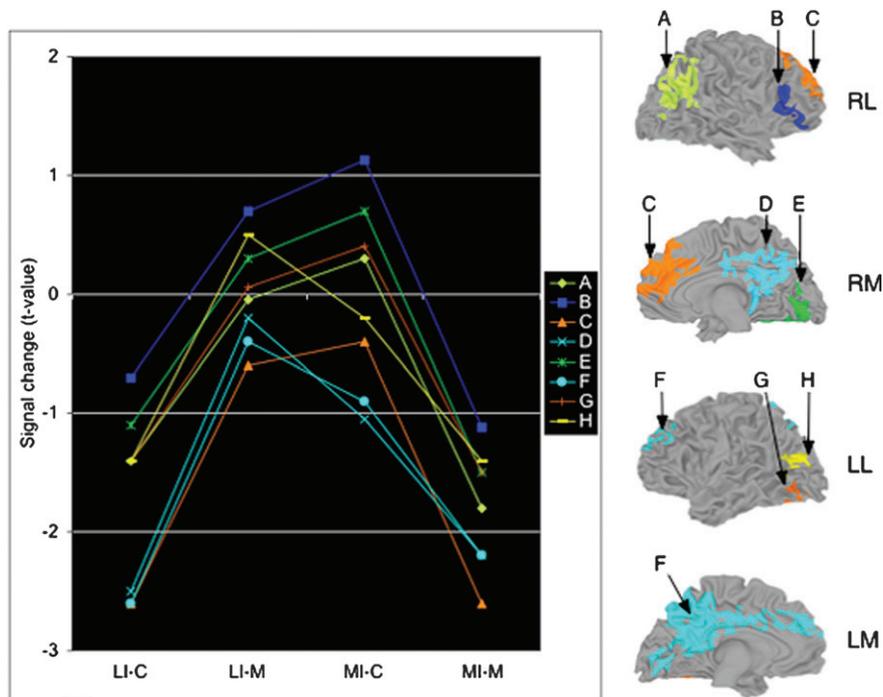


Figure 5. Regional clusters showing an interaction between discourse context and memory performance. Interaction effects (voxel threshold $P < 0.05$, $P < 0.05$ corrected) are marked on lateral and medial views of the right and left hemispheres (RL, RM, LL, LM), with the mean activity in each region for each experimental condition given in the accompanying graph (condition abbreviations as in Fig. 4). The labels of plots in the graph correspond to the subscripts of the marked regions: (A) R. inferior parietal lobule, (B) R. IFG, (C) R. prefrontal cortex, (D) R. posterior midline regions, (E) R. fusiform, (F) L. midline regions, (G) L. fusiform, and (H) L. MTG.

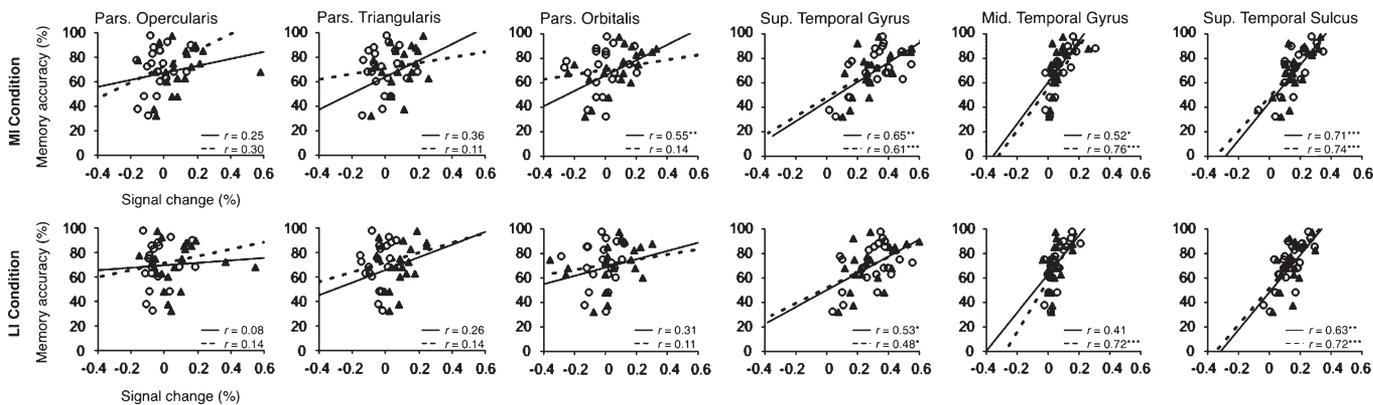


Figure 6. Individual differences in memory accuracy as function of neural activity in inferior frontal and temporal regions. The scatter diagrams plot individual memory accuracy as a function of mean percent signal change in 6 anatomical regions, for the more- and less-informative conditions (MI, LI). For each region, this relation is plotted for the left hemisphere (triangles, unbroken trendline) and right hemisphere homologs (circles, dashed trendline). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

deactivation: Of the 1317 voxels implicated in both effects, only 17 voxels (1.3%) were found in regions demonstrating deactivation (voxel size = 55 mm^3).

Some have argued that finding brain regions implicated in 2 cognitive functions (e.g., IFG implicated in both semantic processing of single words and subsequent memory for those words) suggests that these 2 functions are related (e.g., Gabrieli et al. 1998; Wagner et al. 1998). However, it could also be argued that such overlaps reflect independent functions that share a neural substrate. In the current study, we therefore established 2 independent criteria to assess whether establishing a sentence's contextual meaning and the encoding of that meaning are indeed related: First, if the 2 functions are related

(in the sense that they are cognitively related), then involvement of a region in one function should be *diagnostic* of its involvement in the other. We found that voxels were 4 times more likely to show a subsequent memory effect if they showed a context effect than if they did not show a context effect ($P[\text{mem}|\text{context}] = 0.36$; $P[\text{mem}|\text{not context}] = 0.09$). This association was statistically confirmed by an odds-ratio analysis ($P < 0.001$). Second, if the 2 functions are integrally related, then the degree to which a voxel is sensitive to informativeness in context, as indicated by the magnitude of the context effect for that voxel, should predict the degree to which its activity results in subsequent memory, as indicated by the subsequent memory effect for that voxel. If the functions were

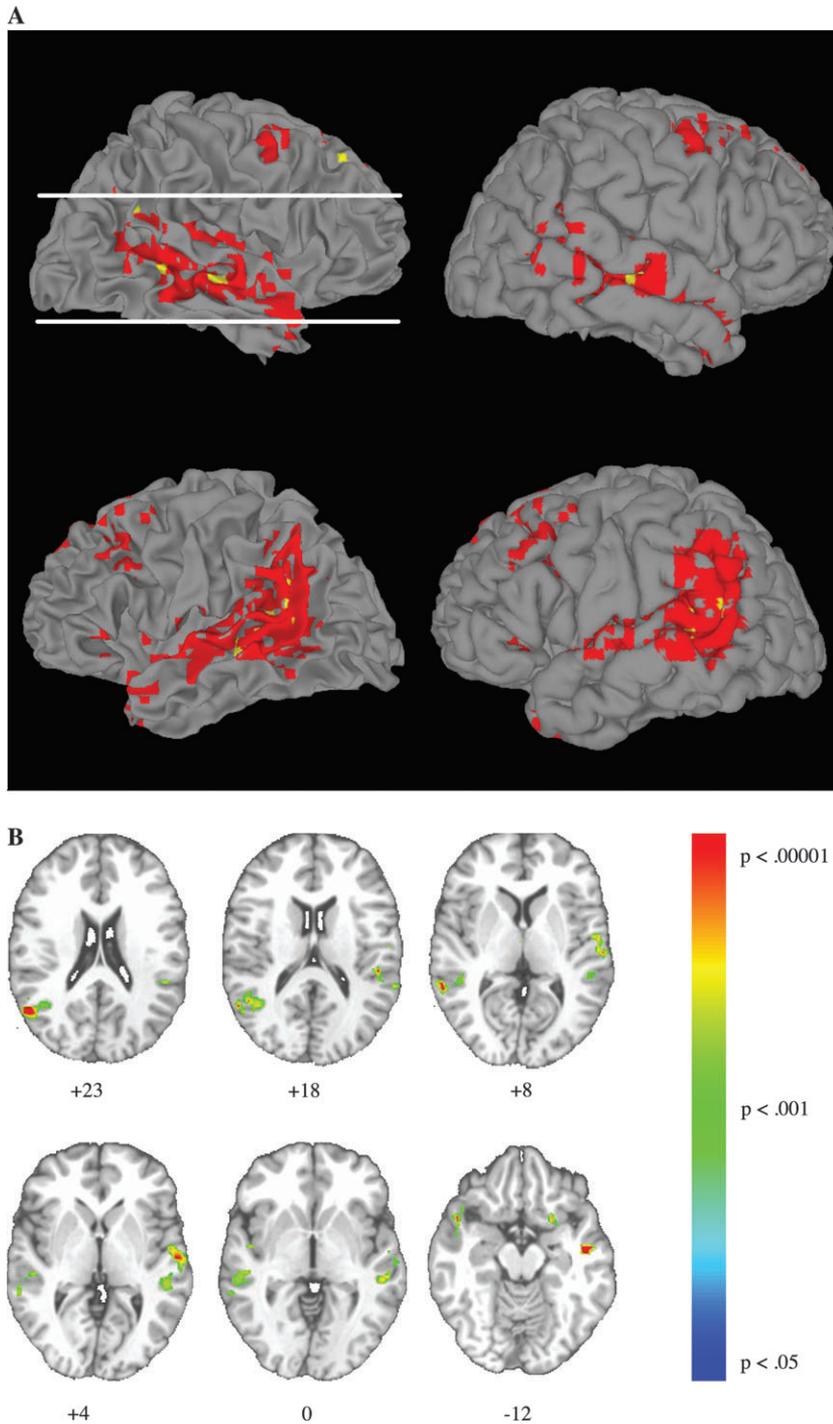


Figure 7. Regions showing both an effect of context and an effect of subsequent memory. (A) Regions showing both effects: Regions in red showed both a reliable effect of context and a reliable effect of subsequent memory (each thresholded at an individual voxel threshold of $P < 0.05$, $P < 0.05$ corrected, see text). Regions in yellow reflect areas independently reliable at an individual voxel threshold of $P < 0.001$ in each contrast (joint probability, $P < 0.000001$). (B) Clusters demonstrating a positive correlation between the magnitude of participants' context and subsequent memory effects (see text). Axial slices are in region bound by the 2 white horizontal lines in (A). Slices presented in radiological convention (left is right).

independent, no association would be expected. To evaluate this association statistically, we conducted a voxel-wise between-participants correlation analysis that assessed for each voxel whether participants that demonstrated a greater context effect also demonstrated a greater subsequent memory effect. We conducted this analysis in those regions where both effects

were reliable (Fig. 7A). This analysis revealed a number of reliable clusters in which all voxels demonstrated a reliable positive correlation between the 2 measures (minimum Pearson's $r > 0.59$ corresponding to individual voxel threshold of $P < 0.005$. FWE $P < 0.05$ corrected; see Fig. 7B). In temporal regions, clusters showing positive correlations were found in

Table 6

Regions showing both context and subsequent memory effects

Region	Volume	x	y	z	BA
L. STG	7268	-53	-45	14	22
R. STG	4756	52	-26	-1	21
L. STG	1282	-50	-25	3	22
R. STG	962	46	6	-16	38
R. MFG	695	42	1	46	6
L. Insula	641	-44	-11	-6	13

Note: Joint probability of individual voxel in mask, $P < 0.0001$; cluster extent of at least 10 voxels (534 mm^3). Center of mass identified by Talairach and Tournoux coordinates.

STG/STS, bilaterally. It is interesting to note that in prefrontal regions one cluster was found with a reliable negative correlation between the measures.

To summarize this analysis, we found that 1) there was substantial overlap between cortical regions showing an effect of informativeness in context and a subsequent memory effect, 2) sensitivity to context was diagnostic of association with subsequent memory, and 3) a number of clusters showed a reliable correlation between the magnitude of the 2 effects.

Discussion

The current fMRI study constituted an initial examination of the cortical networks involved in establishing the information conveyed by sentences in context and those networks involved in the encoding of that information to memory. We found frontal and temporal regions that demonstrated both properties; that is, sensitivity to the informativeness of sentences in context and an involvement in memory encoding. Specifically, these regions showed increased activation for more informative story endings and independently showed greater activity for information that was subsequently remembered. Furthermore, knowing that a brain region was sensitive to the informativeness of sentences was diagnostic of its involvement in memory encoding. However, we also found a number of regions where the association between the semantic processing of sentence meaning and memory encoding was more complex and seemed to involve mediation of below-baseline activation (i.e., modulation of deactivation).

A Link between Semantic Processing and Encoding

A strong relation between semantic processing and encoding was found in dmPFC (most lateral extent, Talairach coordinate $x = 24$ on the right) and also in lateral temporal regions. In these regions, comprehension of sentences in the MI condition was associated with increased activity; independently, increased activity in these regions predicted subsequent memory for content. The role of both these regions in the online integration of content is consistent with prior literature. A number of studies have shown that the dmPFC (BA 8, 9) plays a role in the integration of information: It is associated with the explicit assessment of logical consistency between 2 adjacent statements (Ferstl and von Cramon 2002) or the consistency between a statement and preceding discourse (Ferstl et al. 2005). It is also associated with the integration of information during logical deductive reasoning (e.g., Knauff et al. 2002; Noveck et al. 2004). This region also shows increased activity when sentences are arranged to convey a narrative as compared with when they are presented in random order (e.g., Xu et al. 2005).

Ferstl et al. (2005) suggested that dmPFC activity may be related to increased inference demands when integrating content across sentences, such as when a sentence is less consistent with prior context. But, they also noted that this activity could reflect response uncertainty because in their study participants had to judge each story ending as consistent or inconsistent. Ferstl et al. (2005, p. 733) conclude that "dmPFC activation is not directly caused by external stimulus properties," but elicited by internal evaluation processes. Similarly, Kuperberg et al. (2006) reported increased activity in dmPFC bilaterally for cause-effect sequences of intermediate strength as compared with strongly related or unrelated sequences, but pointed out that given their experimental task (indicating whether each sequence was strongly, intermediately, or not related) this increased activity could reflect greater uncertainty when making judgments for intermediate-strength sequences.

Our findings for dmPFC are important in that they clarify the role of this region in online comprehension because they derive from a passive listening task: First, left dmPFC showed above-baseline activity for both MI and LI conditions (Table 3). Second, right dmPFC showed more activity in the MI condition, and activity in this region also predicted memory for contents (Tables 4 and 5). Also, in the ROI analysis examining brain-behavior correlations we found that activity in this region was predictive of interindividual differences in recognition accuracy for story materials (Supplementary Table S1). These findings indicate that dmPFC is naturally implicated in comprehension of both simple and complex language contents, and furthermore, that this activity can be modulated in the absence of any decision component. Therefore, dmPFC activity reflects processes that are endogenously driven rather than ones driven by exogenous task demands. Taken with prior studies, we suggest that activity in dmPFC during language comprehension involves the natural integration of information, especially when it proceeds nonmonotonically, such as when new information calls for revision of prior information. Its role in language comprehension is likely a specific instantiation of its general function in the organization of information in working memory (Blumenfeld and Ranganath 2006).

Lateral temporal regions have been repeatedly associated with language comprehension (as reviewed in Introduction). In this study, we show for the first time that these regions are strongly associated with the encoding of discourse contents to memory. First, in an anatomically based ROI analysis, we found strong correlations between individual memory performance and neural activity in those regions. Such correlations were particularly strong in the middle third of the right STS and the anterior part of left STS (Supplementary Figure S2), consistent with the purported role of these regions in sentence-level processing (Indefrey and Cutler 2004). Second, when probing for regions that demonstrated greater neural activity for subsequently remembered discourse contents than for subsequently forgotten ones, we found both regional and localized clusters of activation in temporal regions.

Beyond being associated with encoding, temporal regions also exhibited greater activity in the MI condition than in the LI condition. This finding is consistent with the view that lateral temporal regions are involved in integrating discourse information with world knowledge or prior discourse context (e.g., Mazoyer et al. 1993; St George et al. 1999; Xu et al. 2005; Kuperberg et al. 2006; Virtue et al. 2006). Yet, other findings hedge this interpretation: Temporal regions are not sensitive to

whether a statement is true or false (Hagoort et al. 2004), and are not necessarily implicated in determining whether a conclusion necessarily follows from previously read sentences (Ferstl and von Cramon 2002). Thus, we cannot exclude the possibility that the increased activity in the MI condition reflects greater elaboration of the content of single sentences when they are more informative rather than integration per se. A more nuanced pattern of neural activity was found in the right aTL. Most anteriorly, we found an interaction between informativeness and subsequent memory (see Table 5, Fig. 4). For subsequently remembered content, this region showed more activity for more informative sentences, but for subsequently forgotten content it showed the opposite pattern. This suggests that this region plays a particularly important role in encoding of more complex discourse, in that failures in the process are associated with low activity and subsequent forgetting. Immediately posterior to that region, another part of right aTL demonstrated an overlap of the 2 functions (Table 6). Our findings for right aTL are consistent with prior research showing that this region plays a particularly important role in the integration of linguistic information (as opposed to that presented via sounds alone; Humphries et al. 2001), and with research showing that this region is more active during comprehension of less consistent discourse (e.g., Ferstl et al. 2005).

Our findings for left IFG indicate that this region demonstrates increased sensitivity to informativeness along the posterior-anterior axis: The posterior part of the gyrus demonstrated above-baseline activity in both MI and LI conditions, but did not differentiate between them, consistent with its purported role in relatively low-level phonological processes (Poldrack et al. 1999; Gough et al. 2005). A more anterior region demonstrated above-baseline activity for both conditions but additionally showed more activity in the MI condition. Finally, the most anterior aspect of the gyrus showed more activity for the MI condition, but did not show above-baseline in both conditions. This pattern of results strongly supports a functional dissociation between posterior and anterior aspects of left IFG, on which anterior aspects are more specialized for complex semantic processing such as controlled retrieval and selection of semantic content (cf., Badre et al. 2005). Prior studies have revealed similar patterns at the levels of single words (e.g., Wagner et al. 2000; Badre et al. 2005; Gold et al. 2005) or single, context-independent sentences (Rodd et al. 2005). Our converging results, found in a study based on passive listening (rather than on externally cued tasks) are strong validation of this model. Our findings for IFG are also consistent with those of Ferstl et al. (2005), who reported increased anterior IFG activity when story endings were less consistent and with those of Kuperberg et al. (2006) who showed that anterior left IFG is involved in establishing causal links between causes and effects. The subsequent memory effect was found in posterior left IFG, and so the locus of the 2 effects did not overlap. However, the anterior part did demonstrate a reliable correlation between neural activity and individuals' memory accuracy which is consistent with the findings of Casasanto et al. (2002) who used an explicit encoding task. Thus, anterior left IFG seems to play a role in manipulating semantic content during discourse and in the encoding of that content.

Taken together, our results indicate that the dmPFC and lateral temporal regions play an important role in establishing the meaning of sentences in a discourse context. Conjointly, this neural activity determines subsequent memory.

Modulation of Deactivation is Associated with Memory for Discourse Content

We found that activity in midline regions was associated with the informativeness of the content being processed and with its encoding to memory. Our analysis revealed a number of clusters demonstrating differential subsequent memory effects for information given in the MI and LI conditions (Figs 4 and 5 and Table 5). Examining this interaction in relation to the resting baseline revealed a novel and remarkably consistent pattern where greater deactivation was associated with retention in the LI condition, but forgetting in the MI condition.

At least 2 studies have reported subsequent memory effects that were associated with the degree of deactivation in certain brain regions (Clark and Wagner 2003; Daselaar et al. 2004). In these studies, medial prefrontal and posterior midline regions demonstrated deactivation during task performance, while at the same time demonstrating less deactivation for subsequently forgotten items than for subsequently remembered ones. Two additional studies have also reported greater neural activity for subsequently forgotten than subsequently remembered items in these regions but without reference to baseline (Otten and Rugg 2001b; Schott et al. 2006).

The midline regions reported in those studies and the midline regions in which we had found reliable interaction patterns form part of a "default network" that consists of brain regions showing decreased activation (vs. a resting baseline) during the performance of externally induced cognitive tasks, and reliable functional connectivity (for recent reviews, see Fox et al. 2005; Raichle and Gusnard 2005). Its activity has been related to cognitive and emotional functions that are characteristic of the brain's default mode of activity at rest, and its deactivation has been linked to suspension of these functions during performance of exogenously cued tasks (Binder et al. 1999; Gusnard and Raichle 2001; McKiernan et al. 2003, 2006). It has therefore been suggested that greater deactivation may be associated with better retention because it indexes efficient allocation of resources to the cognitive task (Daselaar et al. 2004). Yet, this account cannot explain why in our study greater deactivation was associated with retention in the LI condition but forgetting in the MI condition.

We suggest that the interaction between informativeness and subsequent memory can be understood by considering the relation between successful encoding and the allocation of resources to a given cognitive task. Although we agree with the basic premise underlying the "efficient allocation" account, we propose that efficient allocation of resources is not always synonymous with allocation of resources to the externally presented stimuli. On the one hand, when stimuli are relatively poor in content (i.e., uninformative), efficiency means attending solely to those contents. On the other hand, when stimuli are rich in content (i.e., informative), efficiency can be achieved by elaboration and reflection on those contents via internally generated associations. Previous studies that linked deactivation to subsequent memory used stimuli that were not very rich in content (single words: Clark and Wagner 2003; Daselaar et al. 2004). In such cases, disengaging from the stimuli appears to reflect thoughts unrelated to the task; such disengagement has been associated with less deactivation (McGuire et al. 1996; McKiernan et al. 2006) and can explain why less deactivation has been associated with forgetting (Clark and Wagner 2003; Daselaar et al. 2004). However, stimuli rich in content, such as

natural discourse, clearly lend themselves to elaboration, reflection, and inference generation. Such internal processes may involve disengagement from the external stimuli per se, but are known to improve memory and comprehension (e.g., Craik et al. 1983). We therefore suggest that during comprehension of simple contents (in the LI condition), less deactivation may have been associated with thoughts unrelated to those contents leading to forgetting, whereas during comprehension of complex contents (in the MI condition), decreased deactivation may have been associated with thoughts related to content leading to remembering. On this account then, less deactivation can be associated with either better or worse retention, as a function of the properties of the stimuli in question. The activity patterns for the filler condition in these regions support this account (Fig. 4, dashed lines in regions B, C, and D). As the figure shows, when activity in the filler condition is taken as a reference point, increased activity versus reference tended to be associated with retention in the MI condition, but forgetting in the LI condition. That is, increased activity with respect to nonnarrative sentences could predict either retention or forgetting of discourse content, as a function of its informativeness.

The account we develop here calls for further testing. However, it is consistent with prior results in the literature. Fletcher et al. (2003) report that lateral and medial prefrontal cortex demonstrate a stronger subsequent memory effect for a complex task than for a simple task, an identical pattern to the one we find in medial regions. Mason and Bar (2006) show that generating many associations for a word fragment is associated with less deactivation than generating fewer associations, and Kao et al. (2005) show that increased activity in midline regions during encoding is associated with the internal perception that materials will be subsequently remembered. Our results indicate that the default network is involved in memory encoding, and is sensitive to semantic properties of the encoded materials. These findings suggest that a more complete understanding of the relation between activity in this network and subsequent memory could be gained by employing contents for which internal elaboration could be potentially conducive to the task at hand.

Summary

Sentences represent a central vehicle by which people communicate information. By examining brain activity associated with sentences in different contexts, we identified the biological basis underlying comprehension of sentences that differ in their informativeness. We identified 2 neural networks sensitive to the informativeness of sentences: One network consisted of frontotemporal regions associated with processing of linguistic stimuli and the organization of information in working memory; a second network consisted of brain regions focused on endogenously driven evaluations of content. Both networks were also associated with memory for content. These results indicate that in the brain, establishing the information conveyed by a sentence, that is, its contextually based meaning, involves 2 dissociable networks, both of which are related to processing of sentence meaning and its encoding to memory.

Supplementary Material

Supplementary materials can be found at <http://www.cercor.oxfordjournals.org/>.

Notes

¹In some prior work, sentences that are MI with respect to prior context are referred to as false (vs. true; Singer 2006) or as inconsistent (vs. consistent; Ferstl et al. 2005) even when they are logically consistent with the context. This nomenclature is maintained even when data show that participants often find the “inconsistent” sentences to be consistent with prior context (e.g., 36% of inconsistent trials in Ferstl et al. 2005). According to formal logic and classic theories of semantic information (e.g., Bar-Hillel 1964), 2 statements are said to be inconsistent or contradictory when there is no overlap between the states of affairs denoted by one and those denoted by the other; otherwise, they are consistent. The variable of interest in our study is therefore not consistency, which is a binary categorical property, but the informativeness of the target sentences (cf., Hasson et al. 2005 for a discussion). All statements in the current study were consistent with the prior context but varied in the degree to which they introduced new information.

We thank J. Skipper and G. Josse for their help, and two anonymous reviewers for insightful comments. This work was supported by a grant from the National Institutes of Health: NIH RO1-DC03378. *Conflict of Interest:* None declared.

Address correspondence to Uri Hasson, Human Neuroscience Laboratory, Biological Sciences Division, University of Chicago Hospital Q300, 5841 S. Maryland Avenue MC-2030, Chicago, IL, 60637, USA. Email: uhasson@uchicago.edu.

References

- Albrecht JE, O'Brien EJ. 1993. Updating a mental model: maintaining both local and global coherence. *J Exp Psychol Learn Mem Cogn.* 19(5):1061-1070.
- Argall BD, Saad ZS, Beauchamp MS. 2005. Simplified intersubject averaging on the cortical surface using SUMA. *Hum Brain Mapp.* 27(1):14-27.
- Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron.* 47(6):907-918.
- Bar-Hillel Y. 1964. *Language and information.* Reading (MA): Addison-Wesley.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Rao SM, Cox RW. 1999. Conceptual processing during the conscious resting state: a functional MRI study. *J Cogn Neurosci.* 11(1):80-95.
- Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. 2004. Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci.* 7(3):295-301.
- Blumenfeld RS, Ranganath C. 2006. Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. *J Neurosci.* 26(3):916-925.
- Blumstein SE, Myers EB, Rissman J. 2005. The perception of voice onset time: an fMRI investigation of phonetic category structure. *J Cogn Neurosci.* 17:1353-1366.
- Casasanto DJ, Killgore WD, Maldjian JA, Glosser G, Alsup DC, Cooke AM, Grossman M, Detre JA. 2002. Neural correlates of successful and unsuccessful verbal memory encoding. *Brain Lang.* 80:287-295.
- Clark D, Wagner AD. 2003. Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. *Neuropsychologia.* 41(3):304-317.
- 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.
- Craik FIM, Routh DA, Broadbent DE. 1983. On the transfer from temporary to permanent memory. *Philos Trans R Soc Lond B.* 302: 341-359.
- Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis I: segmentation and surface reconstruction. *Neuroimage.* 9(2): 179-194.
- Daselaar SM, Prince SE, Cabeza R. 2004. When less means more: deactivations during encoding that predict subsequent memory. *Neuroimage.* 23(3):921-927.

- Desai R, Liebenthal E, Possing ET, Waldron E, Binder JR. 2005. Volumetric vs. surface-based alignment for localization of auditory cortex activation. *Neuroimage*. 26(4):1019-1029.
- Duvernoy HM. 1991. The human brain: structure, three-dimensional sectional anatomy and MRI. New York: Springer-Verlag.
- Ferstl EC, Rinck M, von Cramon DY. 2005. Emotional and temporal aspects of situation model processing during text comprehension: an event related study. *J Cogn Neurosci*. 17(5):724-739.
- Ferstl EC, von Cramon DY. 2002. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *Neuroimage* 17:1599-1612.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, et al. 2004. Automatically parcellating the human cerebral cortex. *Cereb Cortex*. 14:11-22.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, et al. 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 33:341-355.
- Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*. 9(2):195-207.
- Fletcher PC, Stephenson CME, Carpenter TA, Donovan T, Bullmore ET. 2003. Regional brain activations predicting subsequent memory success: an event-related fMRI study of the influence of encoding tasks. *Cortex*. 39:1009-1026.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med*. 33:636-647.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Essen DCV, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA*. 102:9673-9678.
- Friederici AD. 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn Sci*. 6:78-84.
- Gabrieli JDE, Poldrack RA, Desmond JE. 1998. The role of left prefrontal cortex in language and memory. *Proc Natl Acad Sci USA*. 95(3):906-913.
- Giraud AL, Truy E, Frackowiak RSJ, Grégoire MC, Pujol JF, Collet L. 2000. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. *Brain*. 123(7):1391-1402.
- Gold BT, Balota DA, Kirchoff BA, Buckner RL. 2005. Common and dissociable activation patterns associated with controlled semantic and phonological processing: evidence from fMRI adaptation. *Cereb Cortex*. 15(9):1438-1450.
- Gough PM, Nobre AC, Devlin JT. 2005. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *J Neurosci*. 25(35):8010-8016.
- Gusnard DA, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci*. 2(10):685-694.
- Hagoort P, Hald L, Bastiaansen M, Petersson KM. 2004. Integration of word meaning and world knowledge in language comprehension. *Science*. 304(5669):438-441.
- Hasson U, Nusbaum HC, Small SL. 2006. Repetition suppression for spoken sentences and the effect of task demands. *J Cogn Neurosci*. 18(12):2013-2029.
- Hasson U, Simmons JP, Todorov A. 2005. Believe it or not: on the possibility of suspending belief. *Psychol Sci*. 16:566-571.
- 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.
- Indefrey P, Cutler A. 2004. Prelexical and lexical processing in listening. In: Gazzaniga MS, Editor. *The cognitive neurosciences III*. Cambridge (MA): MIT Press. p. 759-774.
- Kao YC, Davis ES, Gabrieli JDE. 2005. Neural correlates of actual and predicted memory formation. *Nat Neurosci*. 8(12):1776-1783.
- Keenan JM, Baillet SD, Brown P. 1984. The effects of causal cohesion on comprehension and memory. *J Verbal Learn Verbal Behav*. 23:115-126.
- Knauff M, Mulack T, Kassubek J, Salih HR, Green MW. 2002. Spatial imagery in deductive reasoning: a functional MRI study. *Brain Res Cogn Brain Res*. 13:203-212.
- Kuperberg GR, Lakshmanan BM, Caplan DN, Holcomb PJ. 2006. Making sense of discourse: an fMRI study of causal inferencing across sentences. *Neuroimage*. 33:343-361.
- Long DL, Chong JL. 2001. Comprehension skill and global coherence: a paradoxical picture of poor comprehenders' abilities. *J Exp Psychol Learn Mem Cogn*. 27(6):1424-1429.
- Mason M, Bar M. 2006. The default network and the unconstrained associative brain. Annual Meeting of the Cognitive Neuroscience Society; 2006 Apr 8-11; San Francisco, CA.
- Mason RA, Just MA. 2004. How the brain processes causal inferences in text: a theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychol Sci*. 15(1):1-7.
- Mazoyer BM, 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.
- McGuire P, Paulesu E, Frackowiak R, Frith C. 1996. Brain activity during stimulus independent thought. *Neuroreport*. 7(13):2095-2099.
- McKiernan KA, D'Angelo BR, Kaufman JN, Binder JR. 2006. Interrupting the "stream of consciousness": an fMRI investigation. *Neuroimage*. 29(4):1185-1191.
- McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR. 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J Cogn Neurosci*. 15(3):394-408.
- 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.
- Nichols T, Brett M, Andersson J, Wager T, Poline J-B. 2005. Valid conjunction inference with the minimum statistic. *Neuroimage*. 25:653-660.
- Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*. 15(1):1-25.
- Novick I, Goel V, Smith K. 2004. The neural basis of conditional reasoning with arbitrary content. *Cortex*. 40:613-622.
- O'Brien EJ, Rizzella ML, Albrecht JE, Halleran JG. 1998. Updating a situation model: a memory-based text processing view. *J Exp Psychol Learn Mem Cogn*. 24(5):1200-1210.
- Otten LJ, Henson RN, Rugg MD. 2001. Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. *Brain*. 124:399-412.
- Otten LJ, Henson RN, Rugg MD. 2002. State-related and item-related neural correlates of successful memory encoding. *Nat Neurosci*. 5:1339-1344.
- Otten LJ, Rugg MD. 2001a. Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cereb Cortex*. 11:1150-1160.
- Otten LJ, Rugg MD. 2001b. When more means less: neural activity related to unsuccessful memory encoding. *Curr Biol*. 11(19):1528-1530.
- Paller KA, Wagner AD. 2002. Observing the transformation of experience into memory. *Trends Cogn Sci*. 6:93-102.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*. 10:15-35.
- Raichle ME, Gusnard DA. 2005. Intrinsic brain activity sets the stage for expression of motivated behavior. *J Comp Neurol*. 493(1):167-176.
- Reber PJ, Siwec RM, Gitelman DR, Parrish TB, Mesulam MM, Paller KA. 2002. Neural correlates of successful encoding identified using functional magnetic resonance imaging. *J Neurosci*. 22:9541-9548.
- Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb Cortex*. 15(8):1261-1269.
- Saad ZS, Reynolds RC, Argall B, Japee S, Cox RW. 2004. SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. Proceedings of the 2004 IEEE International Symposium on Biomedical Imaging; 2004 Apr 15-18. IEEE Communications Society. 37. p. 1510-1513.

- Schott BH, Richardson-Klavehn A, Henson RNA, Becker C, Heinze H-J, Duzel E. 2006. Neuroanatomical dissociation of encoding processes related to priming and explicit memory. *J Neurosci.* 26(3):792-800.
- Shannon CE. 1948. A mathematical theory of communication. *Bell Sys Tech J.* 27:379-423.
- Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE. 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J Cogn Neurosci.* 9:648-663.
- Singer M. 2006. Verification of text ideas during reading. *J Mem Lang.* 54:574-591.
- Skipper JI, Nusbaum HC, Small SL. 2005. Listening to talking faces: motor cortical activation during speech perception. *Neuroimage.* 25:76-89.
- St George M, Kutas M, Martinez A, Sereno MI. 1999. Semantic integration in reading: engagement of the right hemisphere during discourse processing. *Brain.* 122:1317-1325.
- Tulving E. 2001. Does memory encoding exist? In: Naveh-Benjamin M, Moscovitch M, Henry L, Roediger I, Editors. *Perspectives on human memory and cognitive aging.* Ann Arbor (MI): Edwards Brothers. p. 6-27.
- Uncapher MR, Rugg MD. 2005. Encoding and the durability of episodic memory: a functional magnetic resonance imaging study. *J Neurosci.* 25(31):7260-7267.
- van Berkum JJA, Hagoort P, Brown CM. 1999. Semantic integration in sentences and discourse: evidence from the N400. *J Cogn Neurosci.* 11(6):657-671.
- van Berkum JJA, Zwitserlood P, Hagoort P, Brown CM. 2003. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cogn Brain Res.* 17(3):701-718.
- Vandenberghe R, Nobre AC, Price CJ. 2002. The response of left temporal cortex to sentences. *J Cogn Neurosci.* 14(4):550-560.
- Virtue S, Haberman J, Clancy Z, Parrish T, Jung-Beeman M. 2006. Neural activity during inference generation. *Cogn Brain Res.* 1084:104-114.
- Vogel K, Bussfeld P, Newen A, Herrmann S, Happe F, Falkai P, Maier W, Shah NJ, Fink GR, Zilles K. 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* 14(1):170-181.
- Wagner AD, Koutstaal W, Maril A, Schacter DL, Buckner RL. 2000. Task-specific repetition priming in left inferior prefrontal cortex. *Cereb Cortex.* 10:1176-1184.
- Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL. 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science.* 281:1188-1191.
- Xu J, Kemeny S, Park G, Frattali C, Braun A. 2005. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage.* 25(3):1002-1015.