

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Neural substrates of narrative comprehension and memory

Tal Yarkoni,^a Nicole K. Speer,^b and Jeffrey M. Zacks^{a,*}

^aDepartment of Psychology, Washington University, St. Louis, MO 63130, USA

^bWICHE Mental Health Program 3035 Center Green Drive, Suite 200, Boulder, CO 80301, USA

Received 18 January 2008; revised 14 March 2008; accepted 29 March 2008
 Available online 15 April 2008

When reading a narrative, comprehension and retention of information benefit considerably from the use of situation models — coherent representations of the characters, locations, and activities described in the text. Here we used functional magnetic resonance imaging (fMRI) to explore the neural mechanisms supporting situation model processing. Participants read blocks of sentences that were either unrelated to one another or formed coherent narratives. A timecourse-based approach was used to identify regions that differentiated narrative-level comprehension from sentence-level comprehension. Most brain regions that showed modulation of activation during narrative-level comprehension were also modulated to a lesser extent during sentence-level comprehension, suggesting a shared reliance on general coherence-building mechanisms. However, tentative evidence was found for narrative-specific activation in dorsomedial prefrontal cortex. Additional analyses identified spatiotemporally distinct neural contributions to situation model processing, with posterior parietal regions supporting situation model construction and frontotemporal regions supporting situation model maintenance. Finally, a set of subsequent memory analyses demonstrated that the boost in comprehension and memory performance observed for coherent materials was attributable to the use of integrative situation models rather than lower-level differences in sentence-level or word-level encoding. These results clarify the functional contributions of distinct brain systems to situation model processing and their mapping onto existing psychological models of narrative comprehension.

© 2008 Published by Elsevier Inc.

Introduction

When readers comprehend texts, they do so by constructing mental representations of the situations described in the texts. These dynamic representations, termed *situation models* (van Dijk and Kintsch, 1983; Zwaan and Radvansky, 1998), integrate readers' prior knowledge about the events, characters, goals, etc. described in a text

with explicitly stated information in order to create a more detailed representation of the text (Zwaan, 1999; Zwaan and Radvansky, 1998). When a reader constructs a situation model, he or she begins by laying a foundation for the mental representation based on the initial information and the reader's prior knowledge (Gernsbacher, 1990; Gernsbacher and Kaschak, 2003). Subsequent information is then mapped onto the developing model, allowing increasingly sophisticated inferences to be made about the nature of the events being represented. Together with word-level and sentence-level processing, the construction of situation models guides reading comprehension and memory (Bransford et al., 1972; Bransford and Johnson, 1972).

A growing number of neuroimaging studies have investigated the neural mechanisms supporting narrative comprehension. Most commonly, narrative-level mechanisms are isolated by contrasting brain activation when reading connected sentences or stories with activation when reading sentences or stories that are unrelated or inconsistent to varying degrees (e.g., Ferstl et al., 2005; Ferstl and von Cramon, 2001, 2002; Fletcher et al., 1995; Giraud, 2000; Hasson et al., 2007; Vokey et al., 2001; Xu et al., 2005). The results of such investigations converge on a distributed network of cortical regions subserving discourse-level comprehension. Many of these regions are known to play a relatively general role in language processing — e.g., areas along the middle and superior temporal gyri and inferior frontal cortex (Binder et al., 1994; Ferstl and von Cramon, 2001; Huettnner et al., 1989; Maguire et al., 1999; Robertson et al., 2000; St George et al., 1999), which show consistent recruitment in a broad range of word-level language tasks (Fiez and Petersen, 1998; Turkeltaub et al., 2002; Vigneau et al., 2006). However, other regions appear to be specifically recruited during comprehension of coherent text, including the anterior temporal lobes (ATL; Ferstl et al., 2007; Mazoyer et al., 1993; Stowe et al., 1998; Stowe et al., 2005) and dorsomedial prefrontal cortex (DMPFC; Hasson et al., 2007; Xu et al., 2005).

Despite the emerging consensus about *which* regions are involved in discourse-level comprehension, several important questions about the relationship between situation model processing and brain activation remain unanswered. First, it is unclear whether narrative comprehension depends on narrative-specific neural mechanisms or more general coherence-building mechanisms that

* Corresponding author. Department of Psychology, Washington University, Campus Box 1125, One Brookings Drive, St. Louis, MO 63130, USA.

E-mail address: jzacks@artsci.wustl.edu (J.M. Zacks).

Available online on ScienceDirect (www.sciencedirect.com).

are also involved in sentence-level comprehension. At least one previous study that directly contrasted narrative-level and sentence-level comprehension found narrative-specific activation in regions such as ATL, posterior MTG, and DMPFC (Xu et al., 2005). However, other studies have observed greater activation in most of these regions when reading or hearing coherent sentences than random word lists (Bottini et al., 1994; Kuperberg et al., 2000; Stowe et al., 1999; Vandenberghe et al., 2002), suggesting that the difference between narrative- and sentence-level comprehension may be quantitative and not qualitative. Alternatively, narrative-specificity may arise at a hemispheric rather than at a regional level. Several fMRI studies have suggested that the right hemisphere is selectively involved in high-level discourse comprehension (Robertson et al., 2000; St George et al., 1999). Again, however, this conclusion is contradicted by other studies that have observed bilateral activation during story reading (Hasson et al., 2007; Maguire et al., 1999), or have found evidence for left hemisphere engagement in discourse processing using lateralized visual-field procedures (Prat et al., 2007). Thus, the precise relationship between narrative-level and sentence-level comprehension remains unclear.

Second, little is known about the temporal dynamics of brain activation during narrative reading. Theoretical models of discourse comprehension posit at least three kinds of processes with temporally distinct contributions to situation model processing: foundation-laying processes that are selectively involved in the initial construction of a situation model (Gernsbacher, 1990); maintenance-related processes involved in keeping information accessible for the duration of a narrative (Garrod and Sanford, 1982; Zwaan and Radvansky, 1998); and updating processes that are transiently invoked whenever the currently-described events are no longer consistent with the global situation model (Morrow et al., 1987; Zwaan et al., 1995; Zwaan and Radvansky, 1998). These distinctions generate predictions that are directly testable at a neurobiological level — e.g., that construction-related activation should occur primarily at the beginning of a narrative; that updating-related activation should vary inversely with narrative coherence; and that the cognitive load associated with situation model maintenance should covary with the complexity of the corresponding narrative. However, only one fMRI study to date has contrasted activation across different epochs of narrative reading (Xu et al., 2005), and the use of an assumed hemodynamic response function in that study precluded detailed investigation of activation timecourses.

Finally, the relationship between brain activation during narrative reading and subsequent memory for the contents of those narratives remains unclear. Behaviorally, the use of situation models aids comprehension and memory of narrative text considerably (Bransford et al., 1972; Gernsbacher et al., 1990); however, the neural mechanisms that support this mnemonic boost have not been fully specified. While many fMRI studies have identified brain activation that predicts memory for individual words (Kirchhoff et al., 2000; Wagner et al., 1998) or sentences (Casasanto et al., 2002), only one previous study has identified memory-predictive activation in the context of coherent narratives (Hasson et al., 2007). The latter study found differential subsequent memory effects as a function of expectancy violation, but did not directly compare subsequent memory effects for coherent and incoherent materials. Moreover, memory for narratives was assessed using only a verbatim recognition test; deeper comprehension of materials was not evaluated. Thus, it is unclear to what extent the subsequent memory effects observed by Hasson et al. (2007) depend on

narrative-level use of situation models vs. word- or sentence-level processes.

The present study investigated these issues in a relatively high powered fMRI experiment ($n=29$; ~20 min of total scan time per experimental condition). Participants read blocks of either coherent short stories (*story* condition) or sets of sentences that had been selected from stories but were then scrambled so they could not be combined into a coherent story (*scrambled* condition). Participants were told that when reading scrambled blocks, they should not try to integrate the sentences, and should simply read and understand the sentences. Several features of the design stand out from previous fMRI studies. First, readers made no overt responses during scanning, allowing us to measure brain activity involved in situation model construction relatively naturalistically. Second, the two experimental conditions differed only in whether or not the text afforded construction of a situation model that spanned multiple sentences. Third, the shape of the hemodynamic response in each reading condition was estimated rather than assumed, enabling identification of complex condition \times time interactions that might not emerge in standard block-level comparisons between conditions. Finally, the relationship between brain activation during reading and behavioral indicators of comprehension and memory was modeled at several different levels, including both verbatim recognition of sentences and deeper comprehension of narrative contents.

Materials and methods

Participants

Twenty-nine participants (ages 18–32, 17 women) volunteered to participate in this study for a cash stipend. Data from an additional 4 participants were discarded due to equipment error. All participants were right-handed native English speakers, with no history of language or reading disorders. Informed consent was obtained in accordance with the guidelines set by the Human Studies Committee at the Washington University School of Medicine.

Materials

The present study used fifty-six scenes from the book *One Boy's Day* (Barker and Wright, 1951). Forty-eight scenes were used in the experiment, and eight scenes were used during the practice session. These scenes described the everyday activities of a seven year-old boy (see Fig. 1 for an example; the full set of stimuli are available online at <http://dcl.wustl.edu/stimuli.html>), and they were used to generate two types of sentence sets. For one type, the sentences from half of the scenes were sampled without replacement to produce sets of unrelated sentences (*scrambled* condition). For the other type, the scenes were left intact to produce sets of related sentences (*story* condition). The sentences in the scrambled and story conditions were counterbalanced across participants by constructing two lists of sentence sets (A and B). In list A, sentences in the first 28 scenes appeared in the story condition and sentences in the remaining 28 scenes appeared in the scrambled condition. In list B, sentences from the first 28 scenes appeared in the scrambled condition, and sentences from the remaining 28 scenes appeared in the story condition. When possible, the positions of the sentences within the scrambled sets were maintained such that sentences that appeared as the first sentences within story sets also appeared as the first sentences

STORY

As Raymond skipped down the aisle toward his desk, he glanced around the room. Whenever his glances met those of the other children, his face lit up in a friendly greeting. At his desk he paused, as if undecided whether to sit down or to find something else to do. He turned and went to the back of the room. He climbed onto a window sill. One of the little girls also climbed up on the window sill. Leaning at least halfway over the sill, he gave the street a brief glance. Just as the teacher came to the back of the room, Raymond dropped to the floor with dexterity. Mrs. Logan frowned at Raymond and at the girl, who was still on the sill. Raymond walked briskly to his desk and sat down.

SCRAMBLED

Mrs. Birch called in a pleasant tone, "Raymond, take a bath and then you can go to bed." Raymond noticed this immediately and asked curiously, "Am I four feet high?" He stood and went toward them in a slow, jogging run. Raymond stopped briefly in front of Sherwin's furniture store. He turned the book and tilted it so that Gregory could see. Mr. Birch cast, and the plug hit the barn. Raymond smiled and glanced down at his plate. As Constance showed the first picture, Raymond laughed. He came out of the bathroom carrying a bottle of hair oil. He sat down in his seat, without bothering to bring his feet in from the aisle. He strolled to the back of the room, turned, and sauntered toward the front of the room.

Yes/no test trials

Mr. Birch was working hard in the garden.
Whenever his glances met those of the other children, his face lit up in a friendly greeting.
He turned the book and tilted it so that Gregory could see.
Mrs. Logan began writing numbers on the board.

Multiple choice test trials

Why did Raymond get off the windowsill?

- A) It was getting cold.
- B) He was getting bored.
- C) The little girl was annoying him.
- D) Mrs. Logan came to the back of the room.

Why did Raymond pause at his desk?

- A) He was trying to decide if he had time to go to the bathroom before class started.
- B) Mrs. Logan had just asked him a question.
- C) He was deciding whether to sit down or find something else to do.
- D) He saw something interesting on the floor.

Fig. 1. Sample stimuli and memory tests used in the current study.

within scrambled sets. The sentences in the scrambled sets were also selected to ensure that there would not be story coherence across sentences. Sentences were assigned to scrambled sets such that the mean number of words per set did not differ across the story (list A: $M=132.33$, $SD=2.35$; list B: $M=132.21$, $SD=1.14$) and scrambled conditions (list A: $M=132.12$, $SD=1.30$; list B: $M=132.21$, $SD=1.82$). The number of sentences per block was held constant across the story (list A: $M=10.71$, $SD=1.30$; list B: $M=10.75$, $SD=68$) and scrambled conditions (list A: $M=10.71$, $SD=0.69$; list B: $M=10.67$, $SD=1.09$).

An LCD projector was used to project stimuli onto a screen positioned at the foot of the scanner, and participants viewed the stimuli through a mirror connected to the head coil. Stimulus presentation and timing were controlled by PsyScope software (Cohen et al., 1993) running on an Apple PowerMac G4 computer (Apple, Cupertino, CA). A PsyScope button box was used to record responses during the behavioral testing session.

Reading task and procedure

Sentences were presented one word at a time in 52-point sans-serif font, and all words were centered on the projection screen. Each word remained on the screen for 200 ms and was followed by a

100 ms delay. An additional 400 ms delay followed the end of a sentence (leading to a 500 ms inter-sentence interval). The scrambled and story sets were presented in 12 runs. Six participants had data from only six ($n=1$), eight ($n=1$), nine ($n=2$), or eleven ($n=2$) runs due to equipment error or participant fatigue. Within each run, sets of sentences (*reading blocks*) were alternated with blocks of fixation, and were preceded by an instruction cue indicating whether the participant should expect a story or a set of scrambled sentences. The instruction cue appeared 4 s prior to the onset of the first word in the reading block, and remained on-screen for 2 s. In each condition, participants were instructed simply to read and understand the sentences for a later memory test.

Each run contained four reading blocks: 2 blocks of scrambled sentences and 2 blocks of story sentences. The order of the blocks was fully counterbalanced across runs, and the scrambled and story sentence sets were randomly assigned to each block for each participant. The reading blocks lasted approximately 49 s (48.71 s to 49.50 s for the scrambled blocks, and 48.38 s to 49.48 s for the story blocks), and the fixation blocks lasted approximately 16 s (depending on the exact length of the reading blocks).

Two different font colors, blue and yellow, were used to eliminate participants' need to remember whether the current block represented the scrambled or story condition. For each participant, all

story sets, including story instruction cues, were presented in one color, and all scrambled sets, including scrambled instruction cues, were presented in the other color. The color-condition mappings were counterbalanced across participants. Participants were given 10–15 min of practice with the task prior to the functional scans (using the eight sets of scrambled and story sentences not used during the 12 functional runs).

Memory test

A memory test followed each of the 12 scanning runs after a 1–2 min delay. The first part of the memory test presented one sentence taken from each reading block, as well as four foil sentences that did not appear at any other point in the experiment, for a total of 24 “old” sentences in each condition and 48 foils. The old and foil sentences were randomly ordered during the memory test, and participants were asked to determine whether each sentence was or was not presented during the previous run. The second part of the memory test presented two four-alternative multiple-choice questions from each story block (48 in total), meant to encourage readers to focus on the context and relationships between sentences in the story blocks. Participants were informed of the memory tests prior to starting the functional runs (see Fig. 1 for an example of the stimuli used in each type of memory test).

Due to malfunctions with the response keys, memory test data could not be collected for 1 participant. Three participants did not complete the experiment (see above), but their partial data (either 20 or 36 out of 48 items) were included in the memory performance analysis. Two participants performed perfectly on recognition test items from the story condition, and therefore could not be included in the subsequent memory analysis due to an absence of within-subject variance. We therefore report memory data for 26 participants.

fMRI data acquisition and preprocessing

Images were acquired on a 3 T Siemens Vision MRI scanner (Erlangen, Germany). A pillow, washcloths, and tape were used to minimize head movement, and headphones and earplugs were used to minimize noise from the scanner. High resolution ($1 \times 1 \times 1.25$ mm) structural images were acquired using a sagittal MP-RAGE T1-weighted sequence. Functional images were acquired using a T2*-weighted asymmetric spin-echo echo-planar sequence, with 32 slices (4.0×4.0 mm in-plane resolution) acquired every 2.048 s (frame). An additional T2-weighted fast turbo spin-echo scan acquired structural data in the same planes as the functional scans. The functional data were corrected for movement using a 6-dimensional affine transformation and warped to a standard stereotactic space (Talairach and Tournoux, 1988). Timing offsets between slices were corrected using cubic-spline interpolation, and slice intensity differences were removed. All data were realigned within and across runs for each participant, and image intensity was normalized for each run to a whole-brain mode value of 1000. The data were spatially smoothed with a Gaussian kernel (full width at half maximum 6.0 mm).

fMRI data analysis

Three sets of fMRI analyses were conducted. First, a data-driven ANOVA analysis identified any brain regions that showed a significant condition \times time interaction, irrespective of the precise nature of the effect. Detailed timecourses of activation in these

regions were then plotted for inspection, and these regions were used as ROIs in all subsequent analyses. Second, focused analyses identified regions in which activation timecourses showed specific temporal profiles, including transient onset activations and linear trends as a function of reading time. Finally, a third set of analyses assessed the relationship between brain activation during reading and subsequent behavioral performance on recognition memory and multiple-choice comprehension tests.

ANOVA analysis

fMRI data were analyzed using a general linear model approach (GLM; Friston et al., 1995). Because we were specifically interested in the temporal dynamics of activation during reading blocks, the data were modeled using a Finite Impulse Response (FIR) approach that allowed independent estimation of each point in the activation timecourse. Each reading block was treated as a single trial that initiated with the instruction cue and persisted for the duration of the block. To allow sufficient time for the hemodynamic response to return to baseline following trial offset (~ 20 s), activation in each condition (story vs. scrambled) was modeled using a FIR set of 35 regressors spanning 71.68 s. All analyses were conducted using in-house software (FIDL).

Regions of interest (ROIs) were identified using a voxel-wise random-effects 2×35 analysis of variance (ANOVA) with reading condition (scrambled vs. story) and time (35 timepoints) as the independent variables. Because we hypothesized that differences between conditions might arise at different points in the timecourse for different regions (e.g., differential effects at trial onset in some regions vs. as a function of reading time in others), ROIs were identified using the condition \times time term of the ANOVA. This approach provides the most powerful omnibus test of interaction effects when the shape of the hemodynamic response is not constrained a priori, and allows a broad range of potential effects to emerge (including standard main effects of condition). F statistics from the ANOVA were converted to z statistics. Due to the high power of the current study, a large portion of the brain showed interactions between condition and time. In order to capture the regions most characteristic of this effect, we therefore used a highly conservative intensity threshold of $z=12$ for each voxel, with a minimum of 9 contiguous voxels in each region. However, because serial autocorrelation in the 35 levels of the time factor violated the independence assumption of the repeated-measures ANOVA, the statistical significance of the time \times condition interaction in each resulting region was assessed using a Greenhouse–Geisser correction to adjust the appropriate degrees of freedom. The corrected F -tests confirmed that interaction effects in all ROIs remained highly significant after adjusting for non-sphericity (all $ps < 0.000001$).

Post-hoc ROI-level tests for a main effect of condition were conducted by extracting the estimated timecourse of activation for each ROI, and then multiplying the timecourse by a contrast formed by convolving a boxcar function with a model hemodynamic response function (Boynton et al., 1996). The resulting magnitude estimates for each region were submitted to paired t -tests contrasting each condition against baseline and the two conditions against one another.

Temporal analyses

In addition to the ANOVA analysis, focused contrasts were tested in order to identify regions that showed transient onset

effects or linear changes in activation as a function of reading duration. Transient onsets were defined as the presence of a significant increase from baseline in either reading condition in at least one of the first 8 frames, followed by a significant decrease from the transient peak in one of the two subsequent frames. The 8-frame window was chosen on the basis of visual inspection of timecourses, which suggested that activation in most regions had reached a stable plateau by the eighth frame. To test for differences between conditions in the magnitude of the transient onset peaks, a paired *t*-test was performed for each region identified by the whole-brain analysis. Peak activation in each region was defined as the maximal amount of activation attained in the first 8 frames of each timecourse.

To identify regions that showed linear trends in activation, we focused specifically on activation during frames 7–26 of the timecourse — the “plateau” period during which the hemodynamic response in most regions had stabilized following trial onset but had not yet begun to decay back to baseline (note that frames 7 and 8 were included in both the onset and linear trend analysis in order to allow for regional variability in the latency of the hemodynamic response). For each voxel, a linear contrast was fitted to these 20 frames for each participant, and the resulting coefficients were tested for a significant linear trend at the second level (i.e., in a random-effects model) using a one-sample *t*-test. An additional constraint for a region to be considered significant was that it had to show a significant overall increase in activation during reading relative to baseline. This constraint was imposed because regions that showed linear changes in activation but no sustained activation during reading were unlikely to be involved in maintenance of narrative-related information, and were not of a priori theoretical interest.

Memory analyses

To identify brain activation that predicted behavioral performance on the memory test following each BOLD run, two different sets of GLMs were constructed. In order to increase power to detect memory effects, all regressors in these GLMs used an assumed response shape. Each regressor was modeled as a boxcar

spanning the duration of the event (see below) convolved with a model hemodynamic response (Boynton et al., 1996). The first set of GLMs included regressors coding for both experimental condition (story vs. scrambled) and within-subject block-by-block recognition memory performance (i.e., hits vs. misses). Recognition memory effects were estimated separately for each reading condition. Moreover, for each participant, recognition memory effects were estimated at both the sentence level and at the block level. Sentence-level regressors spanned only the period during which the target sentence was being presented. For example, if a block contained 10 sentences, and the 7th sentence was later presented as an “old” item during the recognition memory test, the sentence-level regressor would code for the period spanning from the onset of the first word in the 7th sentence until the offset of the final word in the sentence. In contrast, block-level regressors spanned the entire duration of the block from which the target sentence was drawn. This approach enabled separate identification of brain activation that predicted subsequent performance via sentence-level mechanisms vs. block-level mechanisms. We hypothesized that word-level or sentence-level encoding processes should predict subsequent recognition memory performance to a similar extent in both reading conditions, but that narrative-level processes should aid recognition to a greater extent in the story condition, where participants could rely on situation models to aid comprehension and retention of information.

To ensure that the subsequent memory effects identified were not explained by item effects, 4 additional regressors coding for the group-average performance on each item were included in each participant's GLM. Thus, a total of 10 regressors of interest were modeled in each GLM (2 regressors coding for experimental condition; 4 regressors coding for subject-specific recognition memory effects, estimated separately at the sentence level and block level for each experimental condition; and 4 regressors coding for analogous group-average item effects).

The second set of memory GLMs contained all of the regressors included the first set plus two additional regressors coding for block-by-block performance on the multiple-choice comprehension test. One regressor coded for the subject-specific effect, with each story block assigned a value of 0, 1, or 2, reflecting the

Table 1
Regions that showed a condition \times time interaction in the ANOVA analysis

No.	Region	Hemisphere	BA	x	y	z	mm ³
1	Middle/superior temporal gyrus	L	21/22/37	−52	−54	9	13257
2	Inferior frontal gyrus	L	45/46/9	−49	26	6	2781
3	Posterior cerebellum	L		−19	−79	−38	2214
4	Dorsomedial PFC	L	9/8	−11	42	47	3051
5	Left anterior PFC	L	10	−37	49	15	2025
6	Inferior parietal cortex	L	40/7	−42	−58	47	3132
7	Dorsal premotor cortex	L	6	−43	0	50	1485
8	Lingual gyrus	L	17	−10	−95	−2	378
9	Middle /superior temporal gyrus	R	21/22/37	52	−40	5	16470
10	Inferior frontal gyrus	R	45/46	51	28	6	2241
11	Posterior cerebellum	R		23	−78	−34	2808
12	Dorsomedial PFC	R	9	5	53	29	405
13	Right anterior PFC	R	10	38	42	21	5022
14	Inferior parietal cortex	R	40/7	42	−53	48	9963
15	Superior frontal gyrus	R	6/8	10	28	60	297
16	Anterior cingulate cortex	M	32	0	26	35	5076
17	Posterior cingulate cortex	M	23/31/7	0	−35	31	9612
18	Precuneus	M	7/19	1	−76	36	10044

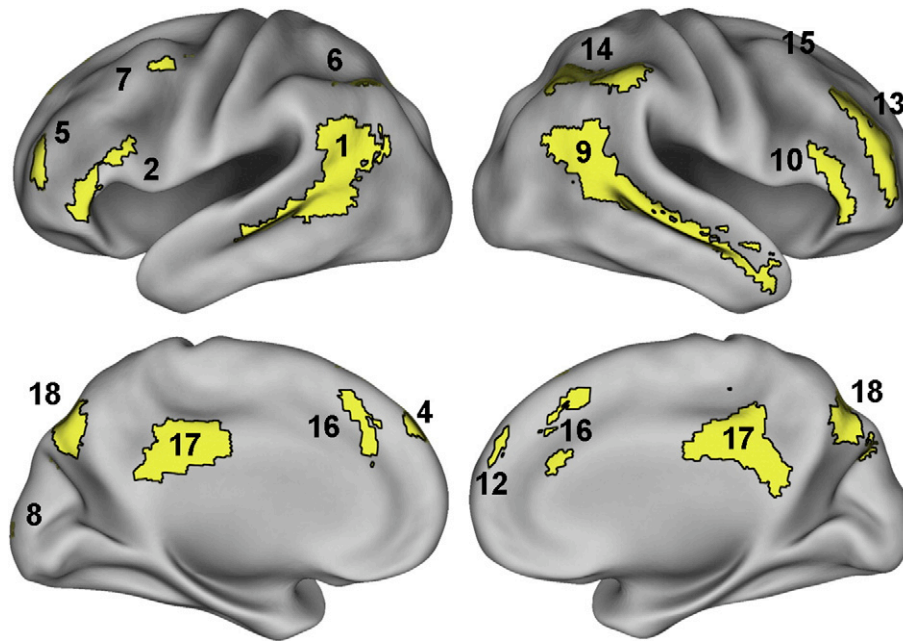


Fig. 2. Regions that showed a significant condition \times time interaction. Number labels correspond to timecourse panels in Fig. 2 and row IDs in Table 1.

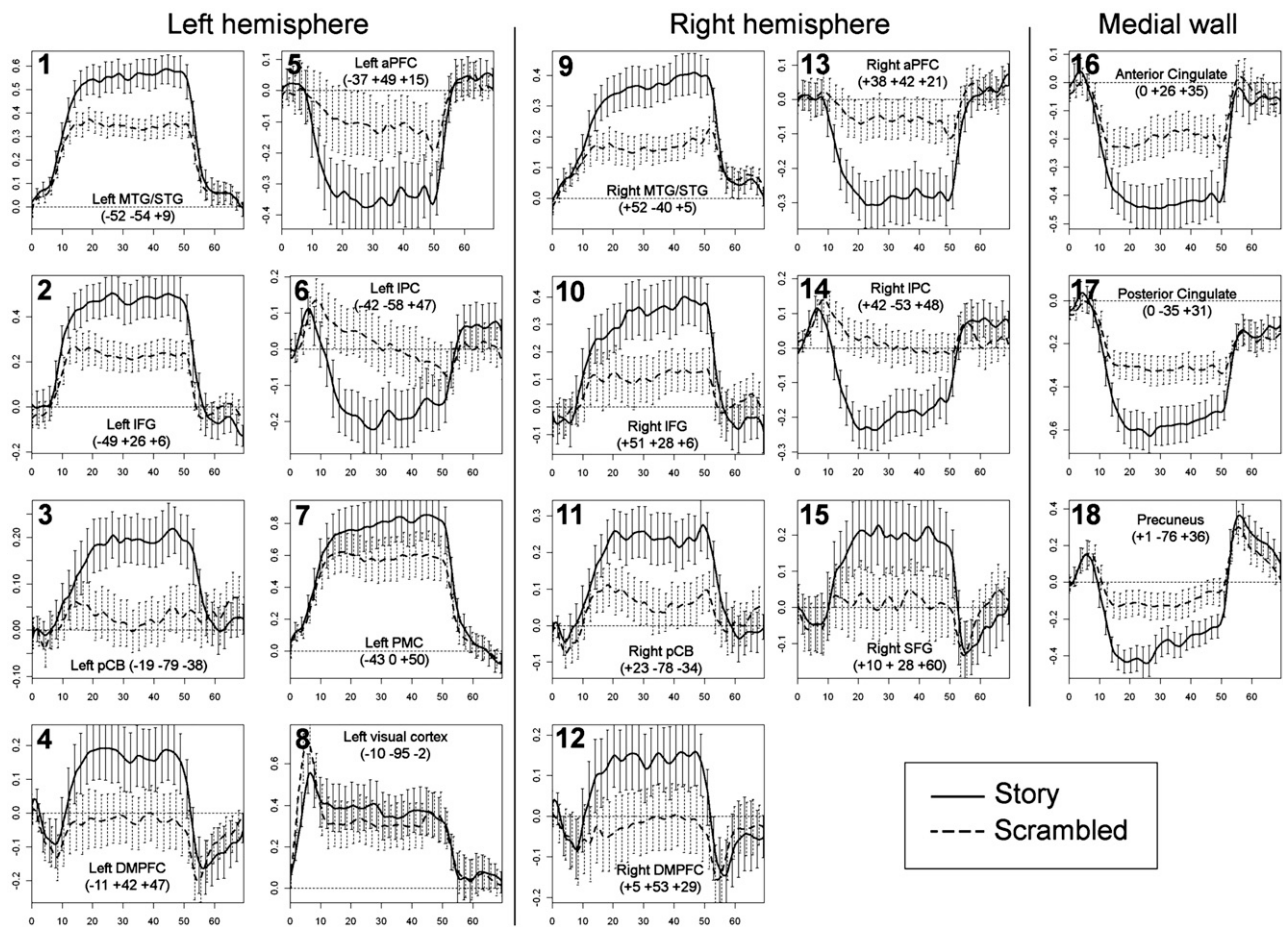


Fig. 3. Timecourses of activation in the story and scrambled conditions. Error bars represent 95% confidence intervals. The x-axis in each panel indicates elapsed time (in seconds) since block onset; the y-axis indicates % change in BOLD activation relative to baseline.

number of correct MC questions the participant answered for that story. The other regressor coded for group-average performance on each item in order to control for item effects. Note that the simultaneous inclusion of regressors for both recognition memory and multiple choice ensured that the resulting estimate for the multiple-choice effect would reflect brain activation associated with multiple-choice performance independently of any recognition memory effect.

Statistical analysis and visualization

In addition to the differing variables of interest described above, all GLMs included 24 regressors coding for effects of no interest (12 coding for differences across each run, and 12 coding for the linear trend within each run). For all ROI-level tests (including tests for onset effects, linear trends, and memory analyses), values for all

voxels within the ROI were first averaged, and the resulting mean was tested using a Type I error protection rate of $p < 0.05$, uncorrected. For all whole-brain analyses except for the condition \times time ANOVA (see above), a voxel-wise (intensity) threshold of $|z| > 3.5$ ($p < 0.0006$) and a cluster-wise (extent) threshold of 9 or more contiguous voxels were used to correct for multiple comparisons. This combination of thresholds has been demonstrated using Monte Carlo simulations to provide an overall whole-brain Type I error rate of $p = 0.05$ given the present level of smoothing (McAvoy et al., 2001). Note that in cases where a test involved a logical conjunction of effects (e.g., the linear trend analysis that required both a significant linear trend and significantly increased activation), each effect was separately corrected for multiple comparisons, following recommendations by Nichols et al. (2005).

For purposes of visualizing activation, statistical maps were mapped onto a three-dimensional representation of the cortical

Table 2
Summary of overall activation, linear trend, and memory effects in the condition \times time ROIs from the ANOVA analysis

Effect	Left hemisphere								Right hemisphere							Medial wall		
	MTG/ STG	IFG	PCB	DMPFC	APFC	IPC	PMC	Lingual. G.	MTG/ STG	IFG	PCB	DMPFC	APFC	IPC	SFG	ACC	PCC	Precuneus
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
<i>Overall activation</i>																		
Story — baseline	++	++	++	++	--	--	++	+	++	++	++	++	--	--	++	--	--	--
Scrambled — baseline	++	++			--		++	+	++	+	+		-			--	--	--
Story — scrambled	++	++	++	++	--	--	++	+	++	++	++	++	--	--	++	--	--	--
<i>Linear trend</i>																		
Linear trend: story — scrambled	++		+				+	-	++	++	+		-					++
Linear trend: scrambled					-	-			+					-				++
Linear trend: story — scrambled	++		++			+	++	-	++	+	++			++				+
<i>Memory</i>																		
Sentence recognition: story	+	+				-	+		+				-	-		-	--	--
Sentence recognition: scrambled	+			-			+		+									
Sentence recognition: story — scrambled																	-	-
Block recognition: story	++	++	+	+	-		++		++	++	+						-	
Block recognition: scrambled																		
Block recognition: story — scrambled	++				-				++	+	+							
Comprehension (story only)	+				-		+				+		-	-			-	-

Numeric column labels parallel ROI labels in Table 1 and Figs. 2–3. Plus and minus signs indicate direction of effect (positive and negative, respectively). Number of characters reflects significance level. One character: $p < 0.05$, uncorrected; two characters: $p < 0.001$, uncorrected.

surface (the PALS atlas; Van Essen, 2005) using Caret software (Van Essen et al., 2001; <http://brainmap.wustl.edu/caret>).

Results

Memory performance

Consistent with the results of previous studies (e.g., Bransford and Johnson, 1972), participants were better able to identify previously-seen sentences when those sentences came from story blocks ($M=0.82$, $SD=0.10$) than from scrambled blocks ($M=0.52$, $SD=0.18$), $t(25)=7.88$, $p<0.001$). Note that although recognition of scrambled sentences was close to 50%, the low hit rate was due to a conservative response strategy rather than to random responding. Participants correctly rejected foils, which were shared across the story and scrambled blocks, at a significant and very high rate ($M=0.88$, $SD=0.10$, $p<0.001$). A signal detection analysis confirmed that d' differed significantly from chance in both the scrambled ($M=1.45$, $SEM=0.13$, $t(25)=3.17$, $p<0.004$) and story ($M=2.45$, $SEM=0.16$, $t(25)=8.69$, $p<0.001$) conditions. Performance on the multiple-choice questions was also well above chance, with participants successfully distinguishing the correct answer from the three foils on an average of 71% of the questions ($SD=0.10$), $t(25)=23.56$, $p<0.001$.

fMRI results

ANOVA results

An initial 2×35 condition \times time ANOVA identified 18 regions that showed a significant condition \times time interaction. These regions included large portions of bilateral temporal cortex extending along the middle and superior temporal gyri (MTG/STG), bilateral inferior

parietal cortex (IPC), bilateral inferior frontal gyrus (IFG) and anterior PFC (APFC), as well as medial regions in the precuneus, anterior and posterior cingulate cortex (ACC/PCC), and dorsomedial prefrontal cortex (DMPFC). Bilateral regions in the posterior cerebellum (PCB) also showed a statistically significant interaction of condition and time. Tables 1 and Figs. 2 and 3 display detailed information and timecourses of activation for each region.

Post-hoc t -tests comparing the overall magnitude of activation between conditions found a significant difference in all 18 ROIs (all $ps<0.001$; Table 2). In all ROIs, this effect reflected greater deflection from baseline in the story condition than in the scrambled condition. More specifically, visual inspection of activation timecourses suggested four distinct patterns of activation (Fig. 3; Table 2). First, in 7 ROIs, activation was significantly greater than baseline in both conditions, and significantly greater in the story condition than in the scrambled condition. These regions included bilateral MTG and IFG, left dorsal premotor cortex (PMC), left lingual gyrus, and right PCB. Second, 5 ROIs, including precuneus, ACC, PCC, and bilateral APFC, showed the opposite pattern: activation decreased significantly from baseline in both conditions, but decreased significantly more in the story condition than the scrambled condition. Third, 4 regions, including bilateral DMPFC/anterior PFC, right superior frontal gyrus (SFG), and left PCB, showed increased activation during the story condition relative to baseline, but no change in the scrambled condition relative to baseline (all $ps>0.3$). For convenience, we collectively refer to the three medial frontal ROIs as DMPFC in subsequent analyses because they were located in adjacent areas of cortex and all showed the same pattern of activation (Fig. 3). Finally, in bilateral inferior parietal cortex, activation decreased from baseline during the story condition, but was not significantly different in the scrambled condition ($ps>0.5$).

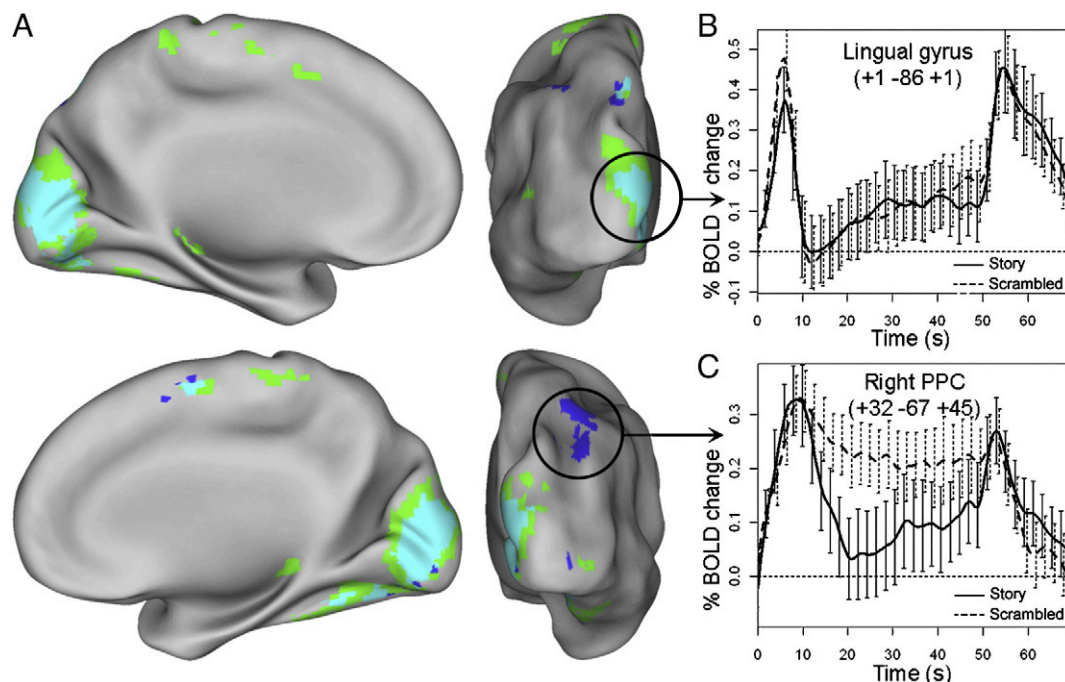


Fig. 4. A. Regions that showed transient positive activation at block onset in the story condition (blue), scrambled condition (green), or both (turquoise). Clockwise from top left: left medial, left posterior, right posterior, and right medial views. B. Common onset effects in visual cortex (activation reflects average of turquoise voxels). C. Selective onset effect in the story condition in PPC.

Temporal dynamics: onset effects

To investigate brain responses associated with initial situation model construction, we searched for regions that showed transient activation increases at the onset of reading blocks. Both ROI-level and whole-brain analyses were conducted. Of the 18 ANOVA ROIs, onset effects were identified in left lingual gyrus, bilateral IPC, and precuneus (Fig. 3). In the left lingual gyrus, the magnitude of the onset peak was significantly greater in the scrambled condition than in the story condition ($t(28)=5.26$, $p<0.001$), suggesting that participants may have been sensitive to the cue information provided by the different word colors in the two conditions. However, transient onset effects in this region were unlikely to reflect narrative processing per se, as similar transients are a standard feature of the BOLD response to sustained stimulation in early visual cortex (Chen et al., 1998; Fox et al., 2005a; Hoge et al., 1999). Onset effects in bilateral IPC and precuneus were also unlikely to reflect foundation-building processes, because all three ROIs showed correspondingly strong offset effects at the end of reading blocks (Fig. 3) as well as sustained deactivation in the story condition. A more plausible explanation is that the onset transients in these regions reflect visual processing of the start cues before the reading blocks, as similar effects have been observed across a range of other tasks (Dosenbach et al., 2006). Thus, no ANOVA ROI showed an onset effect that could plausibly be associated with the initial construction of a situation model.

Whole-brain analysis identified several regions that showed a transient onset effect in at least one reading condition (Fig. 4; Table 3). Onset effects common to both conditions were observed throughout

much of bilateral visual cortex, and the magnitude of the onset peak was generally greater in the scrambled condition than in the story condition, consistent with the ROI-level results. Additionally, several regions showed significant onset effects in only one of the story or scrambled conditions (Fig. 4, blue and green, respectively). In the majority of these cases, visual inspection of activation timecourses suggested that weaker onset effects were present in the other condition as well (onset peak > baseline in both conditions in 21/24 ROIs, $p<0.05$, uncorrected). Moreover, the timecourses of activation in the story and scrambled condition were very similar in most regions (correlation coefficient across timepoints; mean $r=0.86$; Table 3), suggesting that the difference between conditions was generally quantitative and not qualitative. Strikingly, however, bilateral regions in posterior parietal cortex (PPC) showed relatively distinct temporal profiles in the story and scrambled condition (Fig. 4C; Table 3). Onset peaks of similar magnitude were observed in both conditions, but activation subsequently remained elevated only in the scrambled condition.

Temporal dynamics: linear trends

To identify brain activation associated with temporal changes in situation model maintenance load, we searched for brain regions that showed a linear change in activation as a function of reading duration. If the use of situation models facilitates processing of incoming information by decreasing cognitive load or increasing predictability of information, activation in regions related to narrative comprehension should decrease over time. Conversely, if situation models facilitate narrative comprehension by incurring a

Table 3
Regions that showed transient onset effects in either reading condition

Region	Hemisphere	BA	x	y	z	mm ³	Peak diff. (t)	Corr. (r)
<i>Story</i>								
Lingual gyrus	B	17/18/19	1	−86	1	22005	−6.34	0.95
Putamen	L		−17	8	−4	783	1.57	0.85
Precuneus	L	7	−14	−74	39	513	0.32	0.89
Posterior parietal cortex	L	7	−35	−62	48	1512	−1.68	0.36
Medial frontal gyrus	M	6	3	0	54	540	2.57	0.82
Inferior occipital gyrus	R	18	34	−94	−3	675	0.76	1.00
Putamen	R		15	10	−3	567	0.55	0.88
Posterior parietal cortex	R	7	32	−67	45	4509	0.06	0.54
<i>Scrambled</i>								
Lingual gyrus	B	17/18/19	3	−82	2	44658	−6.07	0.95
Fusiform gyrus	L	37	−25	−52	−19	459	−1.03	0.94
Parahippocampal gyrus	L	35	−20	−34	−5	378	−2.15	0.89
Middle occipital gyrus	L	19	−45	−70	6	270	1.50	0.98
Precuneus	L	7	−11	−73	39	324	0.31	0.91
Cingulate gyrus	L	24	−9	2	45	243	−2.27	0.86
Postcentral gyrus	L	5	−22	−37	60	2619	−3.79	0.9
Medial frontal gyrus	L	6	−8	−12	55	324	−1.56	0.87
Medial frontal gyrus	L	6	−10	−24	65	594	−1.7	0.93
Inferior occipital gyrus	R	18	36	−93	−4	972	0.96	1
Parahippocampal gyrus	R	27	19	−33	−2	324	−1.21	0.93
Postcentral gyrus	R	3	41	−18	47	1242	−3.14	0.77
Precuneus	R	7	8	−76	53	405	−1.28	0.9
Medial frontal gyrus	R	6	4	−3	54	459	−0.03	0.84
Paracentral lobule	R	6	5	−33	60	810	−4.11	0.93
Precentral gyrus	R	6	21	−19	64	405	−2.53	0.82

Peak diff. (t): *t*-values of difference between conditions in magnitude of transient onset peaks (story — scrambled). Corr. (r): correlation between story and scrambled activation timecourses across all timepoints. Italics: $p<0.05$, uncorrected; bold: $p<0.001$, uncorrected.

cognitive cost — i.e., by actively maintaining more story-relevant information online as a narrative grows more elaborate — activation in comprehension-related regions should increase over time. However, in either case, activation should show little or no modulation over time in the scrambled condition, because situation models are likely to be of little use when reading disconnected sentences.

ROI-level tests indicated that 8 of the 11 ANOVA regions that showed positive activation in at least one reading condition also showed a significant linear trend in the story condition (smallest $p < 0.02$; Fig. 3; Table 2). In 7 of 8 cases, the slope was positive — i.e., activation increased as a function of reading time. The sole linear decrease occurred in the left lingual gyrus, likely reflecting an adaptation effect rather than narrative-related processing. In all 8 regions, the slope of the linear trend was significantly more positive in the story condition than in the scrambled condition (smallest $p < 0.01$). Only right MTG/STG showed a significant linear trend in the scrambled condition ($p < 0.04$), with activation increasing as a function of reading time.

A complementary whole-brain analysis identified a number of regions that showed linear increases or decreases in activation in at least one reading condition (see Figs. 5 and 6; Tables 4 and 5). Many of the regions that showed linear increases in activation, e.g., bilateral MTG/STG and right ATL, overlapped closely with the ROIs identified based on the condition \times time interaction in the ANOVA. Like those ROIs, the linear trend in these regions was generally present only in the story condition and not in the scrambled condition (Table 4; Fig. 5C). However, several other regions, including bilateral inferotemporal cortex, right PPC, and right dorsal premotor cortex, showed significant linear increases in both conditions, and the slope of these increases generally did not differ across conditions (Table 4; Fig. 5B). Similarly, nearly all regions that showed significant decreases in activation in one condition also showed a corresponding decrease in the other, and again, the slopes rarely differed between conditions (Table 5; Fig. 6). Decreases in activation were restricted primarily to early visual cortex and somatosensory cortex, consistent with the presence of low-level adaptation effects.

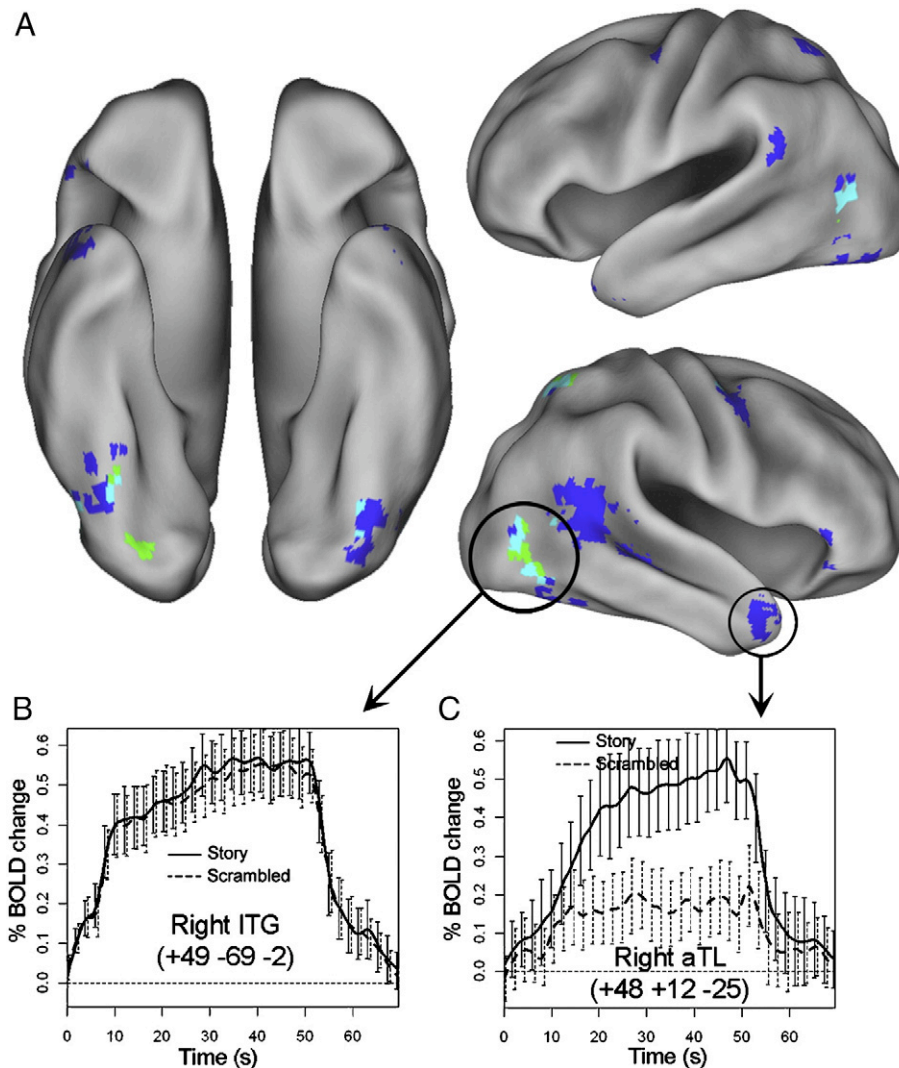


Fig. 5. A. Regions that showed linear increases in activation as a function of reading time in the story condition (blue), scrambled condition (green), or both (turquoise). Clockwise from left: right ventral, left ventral, left lateral, and right lateral views. Timecourses are presented for representative regions showing either a common linear increase in both conditions (B; activation reflects turquoise voxels) or a selective linear increase in the story condition (C).

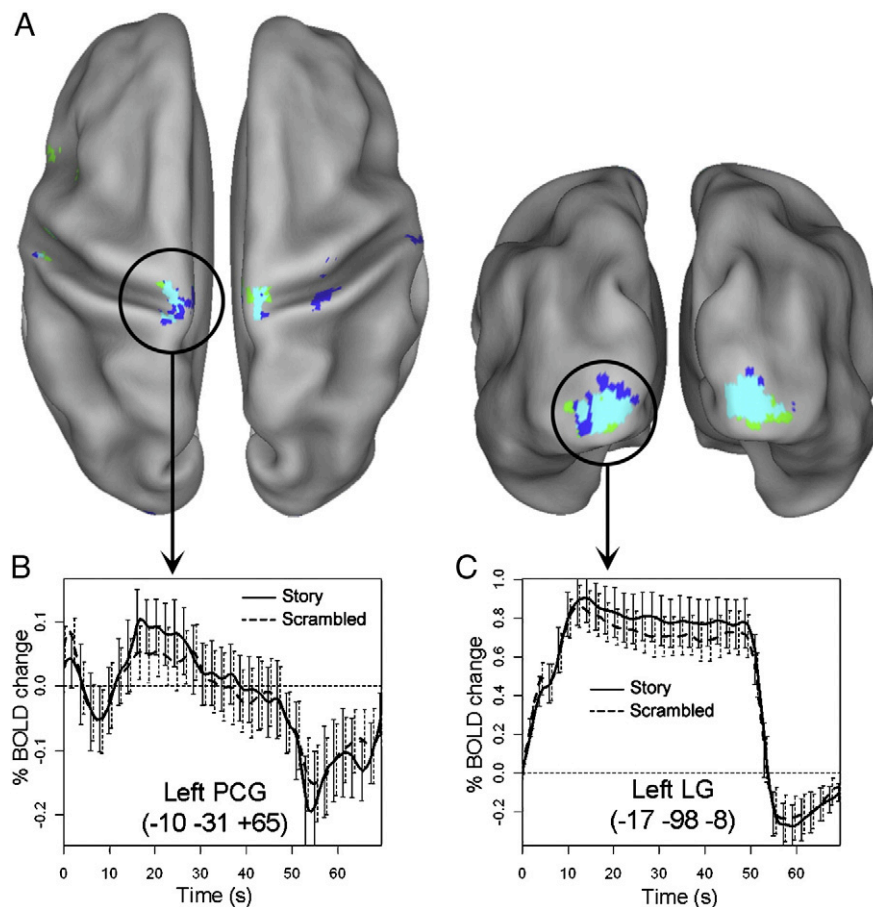


Fig. 6. A. Regions that showed linear decreases in activation as a function of reading time in the story condition (blue), scrambled condition (green), or both (turquoise). From left to right: left dorsal, right dorsal, left posterior, and right posterior cortical views. Timecourses are presented for representative regions in left postcentral gyrus (B) and left lingual gyrus (C); activation reflects turquoise voxels.

Memory and comprehension results: recognition memory

The inclusion of separate sentence recognition and multiple-choice comprehension tests in the present study enabled brain-behavior relationships to be assessed at both a surface processing level (verbatim recognition of individual sentences) and a deeper semantic level (understanding of narrative contents). This approach enabled us to test two competing explanations for the superiority of recognition memory performance in the story condition over the scrambled condition. If superior performance depended on the use of an actively-maintained situation model, recognition memory for sentences should be correlated with differences in activation that sustain over the course of entire reading blocks. If, on the other hand, superior performance in the story condition simply reflected greater task engagement and attentiveness to stimuli, subsequent memory effects should manifest primarily at the sentence level.

ROI-level tests of the 18 condition \times time regions suggested that the mnemonic boost in the story condition primarily reflected block-level mechanisms. A block-level recognition memory effect was observed in 10 ROIs in the story condition, but none in the scrambled condition (Table 2). Moreover, in several of these ROIs, the block-level subsequent memory effect was significantly greater in the story condition than in the scrambled condition. In contrast, sentence-level effects were observed in both conditions: 10 in the story condition and 4 in the scrambled condition (Table 2). Note that in all regions, the direction of subsequent memory effects at

both the sentence and block level mirrored the direction of net activation change. That is, for areas that increased in activation during reading, larger increases were associated with better performance, and for areas that decreased in activation during reading, larger decreases were associated with better performance.

To ensure that regions outside the condition \times time ROIs were not overlooked, we conducted a complimentary whole-brain search for regions that predicted subsequent memory performance at either the story or block level. Two regions (left posterior MTG and right IFG) showed a block-level subsequent memory effect in the story condition (Table 6; Fig. 7); however, both regions were entirely enclosed by the larger ROIs identified by the condition \times time analysis. No region showed a block-level effect in the scrambled condition or a significant difference between the two conditions. In contrast, sentence-level subsequent memory effects were present in 9 regions in the story condition and 2 regions in the scrambled condition (Table 6; Fig. 7). These regions, including portions of ventromedial PFC and a large posterior cingulate region more ventral than the one identified by the condition \times time analysis, showed relatively little overlap with the condition \times time ROIs.

Memory and comprehension results: comprehension

Finally, we investigated whether brain activation during narrative reading predicted comprehension of story contents above

Table 4
Regions that showed linear increases in activation in either or both reading conditions

Region	Hemisphere	BA	x	y	z	mm ³	Story (<i>t</i>)	Scrambled (<i>t</i>)	Story — Scrambled (<i>t</i>)
<i>Story</i>									
Anterior temporal lobe	L	38	−44	14	−30	1539	5.72	2.57	3.11
Inferior temporal gyrus	L	37	−49	−67	0	7911	10.58	4.59	4.99
Inferior frontal gyrus	L	47	−49	22	−7	513	4.46	2.12	2.87
Precentral gyrus	L	6	−39	−6	56	1134	4.27	2.57	4.01
Posterior parietal cortex	L	7	−28	−59	52	432	5.78	2.3	3.51
Anterior temporal lobe	R	38	48	12	−25	2862	6.89	1.24	4.69
Middle temporal gyrus	R	37	53	−54	2	11988	10.85	6.5	5.89
Inferior frontal gyrus	R	47	49	26	−4	729	5.73	2.55	3.16
Middle frontal gyrus	R	6	43	−1	48	1566	6.38	3.02	3.13
Posterior parietal cortex	R	7	25	−59	53	351	5.72	6.07	0.21
<i>Scrambled</i>									
Fusiform gyrus	L	19	−31	−72	−17	351	6.22	5.24	1.04
Inferior temporal gyrus	L		−48	−72	1	1242	10.07	5.57	3.03
Fusiform gyrus	R	37	35	−57	−17	270	4.73	5.17	0.79
Inferior temporal gyrus	R	37	49	−69	−3	2430	7.76	7.41	0.63
Fusiform gyrus	R	19	29	−78	−14	270	3.37	3.67	−1.84
Precentral gyrus	R	6	43	−4	55	243	4.77	4.17	0.61
Precuneus	R	7	25	−58	54	621	5.28	6.5	−0.17
<i>Both</i>									
Fusiform gyrus	L	19	−31	−72	−17	351	6.22	5.24	1.04
Inferior temporal gyrus	L		−48	−72	1	1215	10.11	5.56	3.06
Inferior temporal gyrus	R	37	49	−69	−2	1890	8.52	7.18	0.42
Precuneus	R	7	25	−58	54	297	5.76	6.37	−0.01

The three rightmost columns indicate *t*-values for story vs. baseline, scrambled vs. baseline, and story vs. scrambled activation, respectively. Italics: $p < 0.05$, uncorrected; bold: $p < 0.001$, uncorrected.

and beyond any contribution to verbatim sentence recognition. A new set of GLMs was estimated that included regressors coding for performance on the multiple-choice comprehension test in addition to the existing set of recognition memory regressors. Thus, this analysis identified only comprehension-related activation that was statistically independent of recognition memory-related activation. A significant comprehension effect was observed in 8 ROIs (Table 2). In all cases, the direction of the comprehension effect mirrored the direction of overall activation — better comprehension was associated with increases or decreases of larger magnitude. A complementary whole-brain analysis failed to identify any further regions associated with reading comprehension.

Discussion

The present study identified a distributed network of frontal, temporal, and parietal regions associated with narrative comprehension, broadly replicating the results of previous studies (Ferstl and von Cramon, 2001; Hasson et al., 2007; Mazoyer et al., 1993; Xu et al., 2005). Importantly, however, the present results extend previous findings in several ways. First, the relatively high power of the current study provided a sensitive test of whether narrative-level comprehension recruits qualitatively different neural mechanisms from sentence-level comprehension. No support was found for the notion that narrative-level comprehension depends on the right hemisphere to a greater extent than sentence-level comprehension; however, narrative-specific activations were observed bilaterally in DMPFC. Second, timecourse-based analyses revealed spatiotemporally dissociable patterns of activation that mapped

closely on theoretical distinctions drawn by psychological models of discourse comprehension. Specifically, posterior parietal cortex appeared to be involved in the construction and updating of situation models, whereas perisylvian language areas showed a profile consistent with situation model maintenance. Finally, subsequent memory analyses addressed why it is that coherent stories often lead to better memory than do disconnected sentences. The data strongly suggest that these effects are due to narrative-level use of situation models rather than sentence-level differences in engagement or reading strategy.

Is there hemispheric or regional selectivity for narrative-level comprehension?

Replicating the findings of several previous studies (Ferstl and von Cramon, 2001; Hasson et al., 2007; Mazoyer et al., 1993; Xu et al., 2005), a distributed network of frontal, temporal and parietal brain regions showed significantly greater modulation of activation when reading connected sentences than disconnected sentences in the present study. However, previous results left unclear whether activation in any part of this network is *selective* to narrative-level processing, or if narrative-level and sentence-level comprehension rely on similar coherence-building mechanisms that differ only in the extent to which they are recruited in each condition.

Previous reports have suggested that the right hemisphere is selectively involved in high-level text comprehension (Robertson et al., 2000; St George et al., 1999). In contrast, the present study found no compelling evidence for any lateralization of narrative-level function. This conclusion supports the findings of a recent

Table 5
Regions that showed linear decreases in activation in either or both reading conditions

Region	Hemisphere	BA	<i>x</i>	<i>y</i>	<i>Z</i>	Mm ³	Story (<i>t</i>)	Scrambled (<i>t</i>)	Story — Scrambled (<i>t</i>)
<i>Story</i>									
Lingual gyrus	L	17	−18	−96	−6	3375	−7.88	−5.47	−2.06
Inferior temporal gyrus	L	19	−43	−54	−4	972	−5.59	−4.77	−1.56
Parahippocampal gyrus	L	27	−25	−33	−3	297	−5.21	−5.47	−0.68
Putamen	L		−20	−3	12	621	−6.39	−3.47	−1.83
Insula	L	13	−40	11	19	378	−5.42	−4.53	−0.28
Precentral gyrus	L	4	−52	−13	38	351	−4.68	−4.08	−1.15
Paracentral lobule	L	6	−2	−31	65	2403	−5.83	−4.35	−2.28
Lingual gyrus	R	18	23	−95	−3	3375	−4.61	−4.7	0.21
Precentral gyrus	R	6	60	−7	30	324	−4.73	−3.92	−0.49
Precentral gyrus	R	4	31	−25	61	864	−4.94	−1.76	−1.83
<i>Scrambled</i>									
Lingual gyrus	L	17	−21	−98	−8		−5.93	−5.8	0.69
Inferior temporal gyrus	L	37	−44	−50	−3	1215	−4.48	−5.77	0.15
Inferior frontal gyrus	L	9	−42	12	18	594	−4.32	−4.39	0.77
Precentral gyrus	L	9	−38	4	36	459	0.15	−4.35	4.01
Precentral gyrus	L	4	−50	−13	36	270	−4.29	−4.28	−0.49
Postcentral gyrus	L	4	−10	−31	65	891	−6.22	−4.5	−1.82
Lingual gyrus	R	18	24	−97	−4	4320	−4.18	−4.75	1.36
Paracentral lobule	R	6	7	−28	65	999	−6.4	−4.68	−2.7
<i>Both</i>									
Lingual gyrus	L	17	−17	−98	−8	1890	−7.41	−5.84	−1.16
Fusiform gyrus	L	37	−45	−54	−7	405	−5.32	−4.5	−1.5
Inferior frontal gyrus	L	9	−40	12	19	270	−5.23	−4.5	−0.13
Postcentral gyrus	L	4	−10	−31	65	783	−6.32	−4.53	−1.81
Lingual gyrus	R	18	23	−97	−3	2727	−4.08	−4.56	0.68
Paracentral lobule	R	6	6	−29	65	702	−6.35	−4.71	−2.64

The three rightmost columns indicate *t*-values for story vs. baseline, scrambled vs. baseline, and story vs. scrambled activation, respectively. Italics: $p < 0.05$, uncorrected; bold: $p < 0.001$, uncorrected.

meta-analysis that identified largely bilateral activations associated with the contrast between coherent and incoherent language (Ferstl et al., 2007). Importantly, the present study found bilateral activations not only in the condition \times time analysis that differentiated between coherent and incoherent reading, but also in every other analysis that was conducted. Regions in both hemispheres showed transient onset effects, increased activation as a function of reading time, and showed activation that predicted recognition and comprehension on post-scan tests. The latter findings are particularly informative given that the left and right hemispheres have been associated with local and global perceptual processes, respectively (e.g., Martinez et al., 1997; Robertson et al., 1988; Rossion et al., 2000). If the right hemisphere plays a selective role in extracting the overall gist of a narrative from individual sentences, one might expect right hemisphere regions to show steeper linear increases as a function of reading duration (cf. Xu et al., 2005), or to selectively predict ‘deep’ comprehension rather than verbatim recognition. However, none of these predictions were borne out. Given the ubiquity of bilateral activations in the present study and the relatively small samples used in several previous studies of discourse comprehension (e.g., Robertson et al., 2000; St George et al., 1999), previous reports of selective right-hemisphere involvement in text comprehension may simply reflect a lack of power.

A related question concerns whether there are specific brain regions that are selectively involved in narrative-level comprehension. Previous studies have identified several regions that appear to be recruited to a greater extent during coherent language processing

than incoherent language processing, including DMPFC, ATL and posterior MTG (Ferstl et al., 2007). However, these regions are typically activated during both narrative-level and sentence-level comprehension, suggesting a general involvement in coherence-building rather than narrative-specific comprehension (e.g., Ferstl and von Cramon, 2001; Humphries et al., 2001). Consistent with this interpretation, most regions that showed a significant difference in activation between reading conditions in the present study — including bilateral anterior temporal lobe and MTG — showed significant changes from baseline in both the story and the scrambled condition. The notable exception was DMPFC, in which three ROIs were activated exclusively in the story condition. The latter finding replicates a previous study that identified narrative-selective activation in this region (Xu et al., 2005).

There are at least three plausible explanations for the absence of DMPFC activation in the scrambled condition. First, it may be that DMPFC is in fact engaged in coherence-building at both the sentence and narrative level, but that its activation in the scrambled condition is obscured by a more general pattern of task-related deactivation in this region (cf. Fox et al., 2005b; Raichle et al., 2001). Consistent with this interpretation, inspection of activation timecourses reveals large transient onset and offset deactivations in both conditions in all three DMPFC ROIs (Fig. 3). This pattern appears to be more indicative of countermanding activations and deactivations in DMPFC than of a complete lack of DMPFC involvement in the task (which would presumably produce a flat timecourse).

Table 6

Regions that showed sentence-level or block-level recognition memory effects in either reading condition

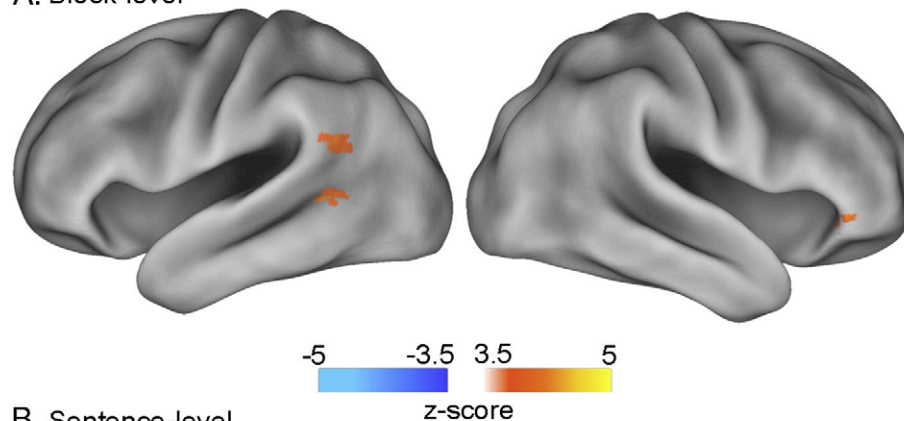
Region	Hemisphere	BA	x	y	z	mm ³	Story (<i>t</i>)	Scrambled (<i>t</i>)	Story — Scrambled (<i>t</i>)
<i>Sentence-level recognition</i>									
Story									
Middle occipital gyrus	L	19	−36	−67	11	74	11.91	0.11	4.55
Insula	L	13	−28	−31	15	16	9.89	1.76	4.00
Supramarginal gyrus	L	40	−52	−48	19	11	6.17	3.83	1.22
Precuneus	L	31	−23	−80	23	13	6.86	1.54	1.84
Inferior parietal lobule	L	40	−62	−34	37	15	5.25	−1.45	4.67
Posterior cingulate cortex	M	23	0	−47	21	344	−7.37	0.29	−4.46
Anterior cingulate cortex	M	32	0	45	10	20	−6.01	−1.90	−2.52
Middle occipital gyrus	R	18	30	−92	13	9	4.22	0.62	1.63
Clastrum	R		29	−15	16	16	5.43	0.41	2.27
Scrambled									
Inferior parietal lobule	L	39	50	−63	44	10	−1.55	−4.71	1.48
Anterior cingulate cortex	R	32	11	39	14	19	−0.70	−7.05	2.29
<i>Block-level recognition</i>									
Story									
Middle temporal gyrus	L	39	−54	−56	9	95	6.82	1.48	4.35
Inferior frontal gyrus	R	47	48	32	−3	17	5.32	0.65	4.22

The three rightmost columns indicate *t*-values for story vs. baseline, scrambled vs. baseline, and story vs. scrambled activation, respectively. Italics: $p < 0.05$, uncorrected; bold: $p < 0.001$, uncorrected.

Second, previous reports of DMPFC activation during sentence-level comprehension could reflect a top-down influence of task instructions (cf. Siebörger et al., 2007). When processing connected sentences, DMPFC activation should occur endogenously and without top-down instruction (Hasson et al., 2007), because readers are naturally motivated to integrate information

across sentence boundaries so as to construct a coherent representation of the narrative. However, when sentences are disconnected and afford no coherent representation, there is little incentive to attempt such integration. In such cases, DMPFC activation may occur only if sentence integration or coherence judgments are explicitly emphasized. Consistent with this view, most studies that

A. Block-level



B. Sentence-level

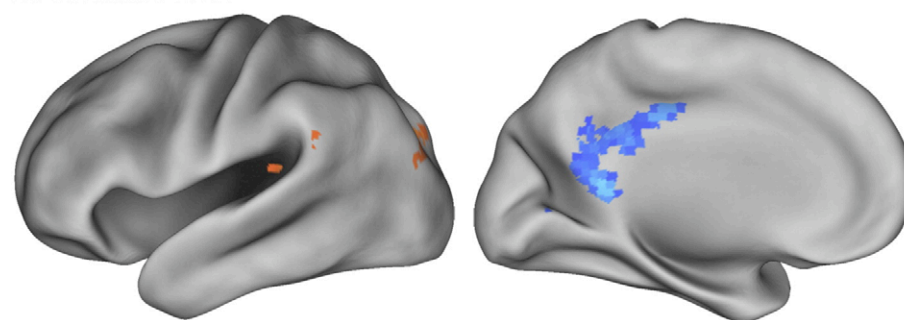


Fig. 7. Regions in which block-level (A) or sentence-level (B) activation predicted subsequent recognition memory for sentences in the story condition.

report DMPFC activation for both coherent and incoherent sentence conditions have required participants to make explicit coherence judgments (Ferstl and von Cramon, 2001, 2002; Kuperberg et al., 2006). Conversely, in both the present study and a previous study that found narrative-specific DMPFC activation (Xu et al., 2005), participants made no overt judgments during scanning and were not instructed to integrate incoherent information. Future studies could test this hypothesis more directly by contrasting activation during reading of incoherent sentences under integration and no-integration instruction conditions.

Finally, it is possible that the absence of DMPFC activation in the scrambled condition reflects a lack of perspective-taking or theory of mind processing. Numerous studies have observed DMPFC activation during tasks that require consideration of agents' beliefs, feelings or intentions (Castelli et al., 2000; Fletcher et al., 1995; Vogeley et al., 2001). In the context of discourse comprehension, Mason and Just (2006) have proposed that DMPFC serves as a "protagonist's perspective" network involved in decoding agents' intentions or goals. On this view, participants might have shown no DMPFC activation in the scrambled condition because it is difficult or impossible to extract coherent representations of agents' motivations and internal states from unrelated sentences. However, a limitation of this view is that it does not explain why coherence judgments can elicit DMPFC activation even in the absence of any theory of mind content (Ferstl and von Cramon, 2002). One possibility is that coherence-related and theory of mind-related functions supported by DMPFC reflect a still more general cognitive function that has not yet been fully characterized.

Dissociable brain systems support the construction and maintenance of situation models

A central aim of the present study was to explore the temporal dynamics of activation during narrative comprehension. Previous fMRI studies of narrative comprehension have focused largely on mean-level differences between coherent and incoherent language conditions (but see e.g., Ferstl et al., 2005; Xu et al., 2005); however, most psychological models of situation model processing explicitly assert that the processing demands associated with narrative comprehension vary over time and reflect distinct cognitive functions. One important functional distinction is between foundation-laying processes associated with the initial construction of a situation model and information-mapping processes involved in subsequent updating of that model based on incoming information (Gernsbacher, 1990). The relative difficulty of constructing a situation model *de novo* is thought to explain why the initial sentence of a story is read more slowly than subsequent sentences (e.g., Gordon et al., 1993; e.g., Haberlandt, 1984). Zwaan and colleagues have similarly proposed an *Event-Indexing* model which assumes that constructive processes are recruited whenever the global situation model is incongruous with the currently processed event and must be updated (Zwaan et al., 1995; Zwaan and Radvansky, 1998). On this view, processing load should be greatest at the onset of a narrative when there is no prior representation of a situation and a completely new spatiotemporal representation must be generated. Thereafter, the amount of updating should vary inversely with the coherence of the events in the narrative.

The present results provide strong support for an Event-Indexing account. Transient increases in activation at block onset were identified in several brain regions. Many of these regions were located in visual and somatomotor areas, and onset effects in these

regions likely reflect basic properties of the BOLD signal rather than narrative-specific processes (Chen et al., 1998; Hoge et al., 1999). However, activation in bilateral PPC showed a response profile remarkably consistent with the predictions of the Event-Indexing model. In both reading conditions, PPC activation showed a large increase from baseline at block onset, presumably reflecting initial construction of a situation model. Subsequently, however, PPC activation decreased markedly in the story condition but remained relatively elevated in the scrambled condition. This divergence of timecourses is precisely what one would expect if model updating depends on the coherence of the events being described. When sentences are connected, each incremental sentence should require relatively few adjustments to the global situation model, because there is little discrepancy between the current and global model. In contrast, when sentences are disconnected, an entirely new situation model must be created for each sentence.

Importantly, the selectivity of updating effects in PPC suggests that the slower reading times associated with initial story sentences are attributable specifically to visuospatial updating processes rather than to a non-specific increase in the amount of cognitive "effort" required to process such sentences. Meta-analyses of fMRI studies implicate PPC activation in a range of executive, working memory, and spatial tasks (Owen et al., 2005; Wager and Smith, 2003; Zacks, 2008) that share as a common denominator the need to manipulate or update actively-maintained visuospatial information (see Wager and Smith, 2003). Presumably, such a function is essential for updating the spatiotemporal representations that situation models consist of (Gernsbacher, 1990; Mason et al., 2006; Morrow et al., 1989; Zwaan and Radvansky, 1998). In contrast, no PPC-like temporal profile was observed in lateral PFC and medial frontal regions generically associated with effortful cognitive processing (Dosenbach et al., 2006; Duncan and Owen, 2000; Yarkoni et al., submitted for publication).

A second issue related to the temporal dynamics of situation model processing concerns maintenance of a situation model over time. Clearly, successful comprehension of a narrative requires that information persist in an accessible form across sentence boundaries. Moreover, the cognitive cost of maintaining a situation model is likely to increase as a narrative grows more complex and the number of events and characters that one must keep track of increases. How and where is such information represented? The present results suggest that situation model maintenance is a distributed process. In left PMC and bilateral MTG, ATL, and IFG, activation was greater when reading connected sentences than disconnected sentences, and increased linearly as a function of reading time primarily or exclusively in the story condition. A parsimonious explanation for these findings is that situation model maintenance occurs relatively automatically during the course of reading. In both the story and the scrambled condition, frontotemporal regions are involved in decoding meaning from text, leading to transient activation of semantic representations. However, in the scrambled condition, the level of activation quickly plateaus, because each incoming sentence makes little or no reference to the characters, settings, and events activated by the previous sentences. In contrast, each sentence in the story condition not only contributes new information, but also associatively reactivates previous representations. As a result, the number of semantic nodes that are active at any given moment is liable to increase over time. Importantly, this focus on passive rather than active maintenance is consistent with psychological models that propose a division of labor between the substantive contents of situation models (thought

to be represented in long-term memory) and the retrieval cues to those contents, which are indexed in WM for efficient integration and updating (Zwaan and Radvansky, 1998).

Of course, this broad functional account should not be taken to imply that regions such as MTG, ATL and IFG play a unitary role in supporting situation model maintenance. Transient reactivation of narrative-related information would be expected to recruit many of the same regions involved in the initial decoding of such information, so the present proposal allows for regional specialization of comprehension-related functions — e.g., that ATL may be involved in high-level propositionalization of narrative contents whereas IFG and MTG are associated with lower-order syntactic and semantic aspects of language processing (for review, see Bookheimer, 2002; Ferstl, 2007; Gernsbacher and Kaschak, 2003).

Interestingly, some brain regions that appeared to be recruited during narrative reading showed no evidence of maintenance-related increases. Specifically, no linear effect of reading time on activation was observed in the three DMPFC regions that showed increased activation in the story condition but not the scrambled condition. Although null results must be interpreted with caution, the widespread presence of highly significant linear trends in other regions suggests that a lack of power is not to blame for the absence of such trends in DMPFC. Rather, these results appear to support the aforementioned view that DMPFC may be involved in strategic, transiently-invoked coherence-building processes (Siebörger et al., 2007) rather than obligatory representation and maintenance of story-related information. Whereas the amount of information contained in a situation model should increase steadily as a narrative unfolds, there is little reason to expect corresponding increases in the need for coherence-building processes. If anything, one might predict the opposite, because the use of situation models should increase the predictability of incoming information (Rinck and Bower, 2000; Zwaan and Radvansky, 1998).

Finally, it is instructive to briefly consider the roles of brain regions that showed significant linear trends in activation but *no* difference between conditions — a pattern identified primarily in visual and somatosensory areas (Figs. 5–6, Tables 3–4). The absence of differences between conditions in these regions is consistent with the notion that their activation reflects non-linguistic processes (in the visual and somatosensory cortex) or sub-sentential processes such as visual word identification (in the inferotemporal cortex; McCandliss et al., 2003; Yarkoni et al., *in press*). Importantly, the presence of such effects demonstrates that between-condition differences in the slope and magnitude of activation observed in frontal and temporal regions cannot be attributed to a general difference in visual attention.

Comprehension and memory for narrative depend on the use of situation models

A final issue addressed in the present study concerned the relationship between brain activation during narrative reading and subsequent comprehension and memory for narrative contents. Specifically, we sought to determine whether the boost in comprehension and memory previously observed for globally coherent narratives (Bransford, 1979; Bransford et al., 1972) reflected narrative-level use of a situation model or lower-level differences in encoding (e.g., deeper processing of individual sentences or words). To our knowledge, only one previous fMRI study has identified brain activation during narrative processing that predicts subsequent recognition memory for sentences. Hasson et al. (2007) identified

memory-predictive activation in a set of frontotemporal regions remarkably similar to those identified in the present study. Additionally, they showed that the relationship between activation in these regions and subsequent memory varied as a function of the informativeness of narrative contents. However, Hasson et al. (2007) did not assess memory for sentences in the scrambled condition, and therefore could not contrast memory for connected and disconnected sentences directly. Moreover, they assessed only verbatim recognition of narratives and not deeper comprehension.

The present results provide strong evidence that the mnemonic boost associated with narrative-level coherence — an increase of 30% in hit rate in the present study — is at least partly mediated by narrative-level mechanisms. Consistent with previous studies demonstrating word-level or sentence-level subsequent memory effects (Casasanto et al., 2002; Davachi and Wagner, 2002; Kirchoff et al., 2000; Wagner et al., 1998), sentence-level activation predicting subsequent recognition memory was identified in both reading conditions. Moreover, activation in bilateral MTG and left premotor cortex — regions commonly associated with word- and sentence-level processing (Fiez and Petersen, 1998; Turkeltaub et al., 2002; Vigneau et al., 2006) — predicted correct sentence recognition in *both* reading conditions. Thus, similar sentence-level mechanisms appear to contribute to recognition memory irrespective of higher-order context. In contrast, block-level subsequent memory effects were observed only in the story condition. Moreover, block-level activation during the story condition predicted not only verbatim recognition of sentences but also performance on a separate multiple-choice comprehension test. These results point to a second, ‘deep’ route mediating memory for sentences drawn from the story condition. That is, participants could identify a sentence drawn from the story condition as an OLD item either because they recognized the precise wording of the sentence or because the events it described were consistent with one of the situation models they constructed while reading the narratives. Such scaffolding was difficult or impossible for disconnected sentences, and readers were therefore forced to rely primarily on sentence-level memory to guide their judgments.

Conclusion

When comprehending narratives, readers construct situation models that aid in the comprehension and retention of processed information. Previous neuroimaging studies identified a distributed set of frontal, temporal, and parietal regions recruited during narrative comprehension; however, the functional contributions of many of these regions to comprehension and memory were left unspecified. The present results help clarify the temporal and mnemonic contributions of distinct brain systems to situation model processing and provide a bridge between neurobiological and psychological models of narrative comprehension.

References

- Barker, R.G., Wright, H.F., 1951. *One Boy's Day: A Specimen Record of Behavior*. Harper.
- Binder, J.R., Rao, S.M., Hammeke, T.A., Yerkin, F.Z., Jesmanowicz, A., Bandettini, P.A., Wong, E.C., Estkowski, L.D., Goldstein, M.D., Houghton, M.D., Hyde, J.S., 1994. Functional magnetic resonance imaging of human auditory cortex. *Ann. Neurol.* 35, 662–672.
- Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev., Neurosci.* 25, 151–188.

- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R.S.J., Frith, C.D., 1994. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain* 117, 1241–1253.
- Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in Human V1. *J. Neurosci.* 16, 4207.
- Bransford, J.D., 1979. *Human Cognition: Learning, Understanding and Remembering*. Wadsworth.
- Bransford, J.D., Johnson, M.K., 1972. Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verbal Learn. Verbal Behav.* 11, 717–726.
- Bransford, J.D., Barclay, J.R., Franks, J.J., 1972. Sentence memory: a constructive versus interpretive approach. *Cogn. Psychol.* 3, 193–209.
- Casasanto, D.J., Killgore, W.D.S., Maldjian, J.A., Glosser, G., Alsop, D.C., Cooke, A.M., Grossman, M., Detre, J.A., 2002. Neural correlates of successful and unsuccessful verbal memory encoding. *Brain Lang.* 80, 287–295.
- Castelli, F., Happé, F., Frith, U., Frith, C., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–325.
- Chen, W.E.I., Zhu, X.H., Kato, T., Andersen, P., Ugurbil, K., 1998. Spatial and temporal differentiation of fMRI BOLD response in primary visual cortex of human brain during sustained visual stimulation. *Magn. Reson. Med.* 39, 520–527.
- Cohen, J.D., MacWhinney, B., Flatt, M., Provost, J., 1993. PsyScope: an interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Beh. Res. Meth. Instrum. Comput.* 25, 257–271.
- Davachi, L., Wagner, A.D., 2002. Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *J. Neurophysiol.* 88, 982–990.
- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for the implementation of task sets. *Neuron* 50, 799–812.
- Duncan, J., Owen, A.M., 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483.
- Ferstl, E.C., 2007. *The Functional Neuroanatomy of Text Comprehension: What's the Story so Far? Higher Level Language Processes in the Brain: Inference and Comprehension Processes*.
- Ferstl, E.C., von Cramon, D.Y., 2001. The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cogn. Brain Res.* 11, 325–340.
- Ferstl, E.C., von Cramon, D.Y., 2002. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *NeuroImage* 17, 1599–1612.
- Ferstl, E.C., Rinck, M., von Cramon, D.Y., 2005. Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. *J. Cogn. Neurosci.* 17, 724–739.
- Ferstl, E.C., Neumann, J., Bogler, C., Yves von Cramon, D., 2007. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* 29 (5), 581–593.
- Fiez, J.A., Petersen, S.E., 1998. Neuroimaging studies of word reading. *Proc. Natl. Acad. Sci. U. S. A.* 95, 914–921.
- Fletcher, P.C., Happé, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S.J., Frith, C.D., 1995. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57, 109–128.
- Fox, M.D., Snyder, A.Z., Barch, D.M., Gusnard, D.A., Raichle, M.E., 2005a. Transient BOLD responses at block transitions. *NeuroImage* 28, 956–966.
- Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005b. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci.* 102, 9673–9678.
- Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.
- Garrod, S.C., Sanford, A.J., 1982. The mental representation of discourse in a focussed memory system: implications for the interpretation of anaphoric noun phrases. *J. Semant. Nijmegen* 1, 21–41.
- Gernsbacher, M.A., 1990. *Language Comprehension as Structure Building*. Lawrence Erlbaum Associates.
- Gernsbacher, M.A., Kaschak, M.P., 2003. Neuroimaging studies of language production and comprehension. *Ann. Rev. Psychol.* 91–115.
- Gernsbacher, M.A., Varner, K.R., Faust, M.E., 1990. Investigating differences in general comprehension skill. *J. Exp. Psychol.: Learn. Memory, Cogn.* 16, 430–445.
- Giraud, A.L., 2000. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. *Brain* 123, 1391–1402.
- Gordon, P.C., Grosz, B.J., Gilliom, L.A., 1993. Pronouns, names, and the centering of attention in discourse. *Cogn. Sci.* 17, 311–347.
- Haberlandt, K., 1984. Components of sentence and word reading times. *New Methods Read. Compr. Res.* 219–251.
- Hasson, U., Nusbaum, H.C., Small, S.L., 2007. Brain networks subserving the extraction of sentence information and its encoding to memory. *Cereb. Cortex* 17 (12), 2899–2913.
- Hoge, R.D., Atkinson, J., Gill, B., Crelier, G.R., Marrett, S., Pike, G.B., 1999. Stimulus-dependent BOLD and perfusion dynamics in human V1. *NeuroImage* 9, 573–585.
- Huettner, M.I., Rosenthal, B.L., Hynd, G.W., 1989. Regional cerebral blood flow (rCBF) in normal readers: bilateral activation with narrative text. *Arch. Clin. Neuropsychol.* 4, 71–78.
- Humphries, C., Willard, K., Buchsbaum, B., Hickok, G., 2001. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *NeuroReport* 12, 1749–1752.
- Kirchhoff, B.A., Wagner, A.D., Maril, A., Stern, C.E., 2000. Prefrontal–temporal circuitry for episodic encoding and subsequent memory. *Journal of Neuroscience* 20, 6173.
- Kuperberg, G.R., McGuire, P.K., Bullmore, E.T., Brammer, M.J., Rabe-Hesketh, S., Wright, I.C., Lythgoe, D.J., Williams, S.C.R., David, A.S., 2000. Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. *J. Cogn. Neurosci.* 12, 321–341.
- Kuperberg, G.R., Lakshmanan, B.M., Caplan, D.N., Holcomb, P.J., 2006. Making sense of discourse: an fMRI study of causal inferencing across sentences. *Neuroimage* 33, 343–361.
- Maguire, E.A., Frith, C.D., Morris, R.G.M., 1999. The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain* 122, 1839–1850.
- Martinez, A., Moses, P., Frank, L., Buxton, R., Wong, E., Stiles, J., 1997. Hemispheric asymmetries in global and local processing: evidence from fMRI. *NeuroReport* 8, 1685–1689.
- Mason, R.A., Just, M.A., Pittsburgh, P.A., 2006. Neuroimaging contributions to the understanding of discourse processes. *Handbook Psycholinguist.* 765–799.
- Mazoyer, B.M., 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, 467–479.
- McAvoy, M.P., Ollinger, J.M., Buckner, R.L., 2001. Cluster size thresholds for assessment of significant activation in fMRI. *NeuroImage* 13, S198.
- McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299.
- Morrow, D.G., Bower, G.H., Greenspan, S.L., 1989. Updating situation models during narrative comprehension. *J. Mem. Lang.* 28, 292–312.
- Morrow, D.G., Greenspan, S.L., Bower, G.H., 1987. Accessibility and situation models in narrative comprehension. *J. Mem. Lang.* 26, 165–187.
- Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. *NeuroImage* 25, 653–660.
- Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59.

- Prat, C.S., Long, D.L., Baynes, K., F., 2007. The representation of discourse in the two hemispheres: an individual differences investigation. *Brain Lang.* 100, 283–294.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98, 676–682.
- Rinck, M., Bower, G.H., 2000. Temporal and spatial distance in situation models. *Mem. Cogn.* 28, 1310–1320.
- Robertson, L.C., Lamb, M.R., Knight, R.T., 1988. Effects of lesions of temporal–parietal junction on perceptual and attentional processing in humans. *J. Neurosci.* 8, 3757.
- Robertson, D.A., Gernsbacher, M.A., Guidotti, S.J., Robertson, R.R., Irwin, W., Mock, B.J., Campana, M.E., 2000. Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychol. Sci.* 11, 255–260.
- Rossion, B., Dricot, L., Devolder, A., Bodart, J.M., Crommelinck, M., de Gelder, B., Zoontjes, R., 2000. Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *J. Cogn. Neurosci.* 12, 793–802.
- Siebert, F.T., Ferstl, E.C., von Cramon, D.Y., 2007. Making sense of nonsense: an fMRI study of task induced inference processes during discourse comprehension. *Brain Research* 1166, 77–91.
- St George, M., Kutas, M., Martinez, A., Sereno, M.I., 1999. Semantic integration in reading: engagement of the right hemisphere during discourse processing. *Brain* 122, 1317–1325.
- Stowe, L.A., Broere, C.A.J., Paans, A.M., Wijers, A.A., Mulder, G., Vaalburg, W., Zwarts, F., 1998. Localizing components of a complex task: sentence processing and working memory. *NeuroReport* 9, 2995–2999.
- Stowe, L.A., Paans, A.M., Wijers, A.A., Zwarts, F., Mulder, G., Vaalburg, W., 1999. Sentence comprehension and word repetition: a positron emission tomography investigation. *Psychophysiology* 36, 786–801.
- Stowe, L.A., Haverkort, M., Zwarts, F., 2005. Rethinking the neurological basis of language. *Lingua* 115, 997–1042.
- Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system — an approach to cerebral imaging. Thieme Medical Publishers, New York.
- Turkeltaub, P.E., Eden, G.F., Jones, K.M., Zeffiro, T.A., 2002. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* 16, 765–780.
- van Dijk, T.A., Kintsch, W., 1983. Strategies of discourse comprehension. Academic Press, San Diego, CA.
- Van Essen, D.C., 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *NeuroImage* 28, 635–662.
- Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., Anderson, C.H., 2001. An integrated software suite for surface-based analyses of cerebral cortex. *J. Am. Med. Inform. Assoc.* 8, 443.
- Vandenberghe, R., Nobre, A.C., Price, C.J., 2002. The response of left temporal cortex to sentences. *J. Cogn. Neurosci.* 14, 550–560.
- Vigneau, M., Beaucousin, V., Herve, P.Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*. 30 (4), 1414–1432.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N.J., Fink, G.R., Zilles, K., 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* 14, 170–181.
- Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect Behav. Neurosci.* 3, 255–274.
- Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., Buckner, R.L., 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, 1002–1015.
- Yarkoni, T., Speer, N.K., Balota, D.A., McAvoy, M.P., Zacks, J.M., in press. Pictures of a thousand words: investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *NeuroImage*.
- Yarkoni, T., Barch, D.M., Gray, J.A., Conturo, T., Braver, T.S., submitted for publication. Neural correlates of trial-by-trial variability in reaction time: a multi-study fMRI analysis.
- Zacks, J.M., 2008. Neuroimaging studies of mental rotation: a meta-analysis and review. *J. Cogn. Neurosci.* 20, 1–19.
- Zwaan, R.A., 1999. Situation models: the mental leap into imagined worlds. *Curr. Dir. Psychol. Sci.* 8, 15–18.
- Zwaan, R.A., Langston, M.C., Graesser, A.C., 1995. The construction of situation models in narrative comprehension: an event-indexing model. *Psychol. Sci.* 6, 292–297.
- Zwaan, R.A., Radvansky, G.A., 1998. Situation models in language comprehension and memory. *Psychol. Bull.* 123, 162–185.